THE ROLE OF BACTERIOPHAGE IN PROVIDING BACTERIAL DIVERSITY IN ENGINEERED AND NATURAL ECOSYSTEMS TO PROVIDE SYSTEM RESILIENCY

by

Seyed Amir Hossein Mohaghegh Motlagh

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Civil and Environmental Engineering

The University of Utah

August 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Seyed Amir Hossein Mohaghegh Motlagh**

has been approved by the following supervisory committee members:

Ramesh K. Goel	, Chair	**12/16/2015** Date Approved
P.K. Andy Hong	, Member	**12/16/2015** Date Approved
Luther McDonald	, Member	**12/16/2015** Date Approved
Milind Deo	, Member	**12/16/2015** Date Approved
Edward Trujillo	, Member	**12/16/2015** Date Approved

and by **Michael E. Barber** , Chair/Dean of

the Department/College/School of **Civil and Environmental Engineering**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Bacteriophages as the most abundant biological entities on the planet play a significant role in microbial population dynamics in various ecosystems. The potential of bacteriophages as a driving force in evolution of microbial communities through controlling the bacterial population, naturally selecting phage-resistant bacteria, and facilitating horizontal gene transfer have been studied. However, few studies have demonstrated the effect of phages on the microbial communities in different natural ecosystems, biological nutrient removal reactors, and hypersaline environment. In this study, the role of bacteriophages in functional gene transfer and how they affect different nutrient cycles and bacterial diversity and population in lab-scale and natural ecosystems were investigated.

The research was accomplished through studying the bacteriophage population, diversity, and their role in bacterial infection and subsequent alteration in bacterial population and diversity. The ecosystems studied in this study included lab-scale biological phosphorus removal and hypersaline Great Salt Lake as engineered and natural models for understanding the phage-host interaction. The biomass and sediment samples were collected from the lab-scale bioreactor and deep brine layer in Great Salt Lake and subjected to various environmental stress factors to understand the role of bacteriophage and prophage induction in bacterial diversity. In addition, the sediment sample from the

Great Salt Lake was analyzed with metagenomics approach.

The evaluation of prophage induction showed that various environmental stress factors including nutrients, heavy metals, toxic chemical, and antibiotic can induce phages integrated onto bacterial genomes (i.e. prophages), resulting in a decrease of the bacterial population involved in different nutrient cycles. Analyzing the viral and bacterial metagenomes explored the GC content, oligonucleotide and k-mer profile, genetic homology, CRISPRs, and prophage network. Our in-depth metagenomics analysis identified phage and bacteria communities comprehensively and proved the role of bacteriophages in defining the bacterial community population, diversity, and their effects on various nutrient cycles. Identification of bacteriophage diversity, population, and their functional genes using metagenomics approach in this study will shed light on the bacterial and viral diversity in Great Salt Lake and this information will be helpful in constructing metabolic models to better study the microbial interaction in various hypersaline ecosystems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my most sincere gratitude to my advisor and mentor, Dr. Ramesh Goel, for his motivational guidance, support, and patience throughout this research. Also, I would like to appreciate my committee members, Dr. Andy Hong, Dr. Luther McDonald, Dr. Milind Deo, and Dr. Edward Trujillo, for valuable discussions and their guidance.

I would also like to thank all of our collaborators who were involved in this research. I appreciate the advice and assitance from Dr. Sherwood Casjens at the Department of Pathology, Dr. Bas Dutilh at the Department of Theoretical Biology of Utrecht University, Dr. Bryan Dalley at the Huntsman Cancer Institute, Derek Warner at the DNA Sequencing Facility, Nancy Chandler at the Electron Microscopy Facility, and Ryan Rowland at the U.S. Geological Survey. In addition, I am very grateful to my lab-mates, particularly Ananda Bhattacharjee and Shireen Meher Kotay, for the inspirational discussions. Also, particular thanks go to MJ Ameli, a great friend and office-mate.

This dissertation would not have been possible without the love, support, and encouragement I received from my family. I would like to appreciate particularly from my lovely wife, Negar, that undoubtedly, it was not possible to finish this journey without her love, support, understanding, and patience. Finally, I would like to thank all those who have helped put this dissertation in its final form.

CHAPTER 1

INTRODUCTION

1.1. Overview

Several chemical, physical, and biological processes related to nutrient, carbon, and metal cycling can occur simultaneously in natural ecosystems such as rivers, lakes, and wetlands. These processes are beneficial in removing excess nutrients and pollutants (Fisher and Acreman, 2004; Jasper et al., 2013), transforming heavy metals (Turner et al., 2000), or enhancing nutrient consumption in plant biomass (Kobayashi et al., 2013).

Sustainable functioning of any natural ecosystem requires maintaining the resilience of the system and its ability to cope with different environmental stresses, shocks, and enhance capabilities that allow adaptation (Turner et al., 2003). The resilience of such a system is directly correlated with the microbial diversity present in the environment, as these organisms are actively involved in contaminant mobilization through the different strata. Microbial processes in the sediments of a natural ecosystem play an important role in water purification through nutrient cycling (Sims et al., 2012; Wang et al., 2012), carbon cycling (Jasser et al., 2009; Li et al., 2011), mercury transformation (Chavan et al., 2007; Mitchell et al., 2009; Ramasamy et al., 2012), sulfate reduction (Bahr et al., 2005; Whitmire and Hamilton, 2005; Faulwetter et al., 2013), and other biological processes. As a result, the microbial community plays a

major role in overall functioning of various natural ecosystems (Danovaro and Pusceddu, 2007; Reed and Hicks, 2011). The common thing among all of these biological processes is that these all are prokaryotes-mediated. In other words, natural ecosystems have bacterial diversity which provides environmental sustainability to these ecosystems and enables biologically and chemically disparate microzones to co-occur, facilitating diverse ecological processes in a small volume. Along with prokaryotes, other entities which also co-exist in natural ecosystems are bacteriophages, i.e. viruses that infect only bacteria. Bacteriophages are the most numerous and diverse biological entities in aquatic environment and on earth (Suttle, 2007). Although bacteriophages do not directly participate in any metabolic activity in ecosystems, they can affect the resiliency of any ecosystem by directly affecting the prokaryotic community through their well-known infection cycles. Viral infection and lysis are among the biotic bottom-up control factors that have significant stimulatory and/or inhibitory effects on biodiversity of prokaryotes, microbial biomass, and their activity in the aquatic system. This is a common saying that bacteriophages exist wherever bacteria are living. In fact, bacteriophage number in many cases outcompete the bacterial number, which suggests that there must be significant phage-bacteria interactions in natural ecosystems.

Interactions between bacteriophages and prokaryotes have rarely been examined in natural ecosystems and the role of viruses in regulating microbial communities in such an ecosystem is unclear (Jackson and Jackson, 2008). Bacteriophages have been studied extensively in marine environments and more recently in freshwater, but very few studies have examined viruses in hypersaline ecosystems and shallow water such as wetlands. A quick search using major published paper databases enabled us with less than 30 papers on viruses in wetland ecosystems, of which most of the papers dealt with either transport

of viruses with coliphages as model viruses or human viruses.

As mentioned in the previous section, although bacterial-phage interaction plays a significant role in various natural ecosystems, very few studies focused on this interaction and its effects on bacterial diversity, their population, and the system resiliency to these changes. In addition, to best of our knowledge, there is no study conducted to understand the diversity of bacteriophages in a hypersaline environment and their role in affecting the nutrient removal and metal transformation that is happening in this ecosystem.

This research was conducted to better understand the role of bacteriophages and their diversity and how these affect bacterial community structure in various ecosystems including a lab-scale biological phosphorus removal and hypersaline ecosystem. To achieve the objectives of this research, three hypotheses were designed as follows: (i) environmental stress factors that various ecosystems face cause bacterial mortality through prophage induction; (ii) lytic phages co-exist with bacterial communities in ecosystems' biomass and sediments and can affect bacterial diversity by infecting the dominant species; (iii) bacteriophage-mediated gene transfer is leading to diversification of microbial communities in various ecosystems.

1.2. References

Bahr, M.; Crump, B. C.; Klepac-Ceraj, V.; Teske, A.; Sogin, M. L.; Hobbie, J. E. Molecular characterization of sulfate-reducing bacteria in a New England salt marsh. *Environ. Microbiol*. **2005**, *7*, 1175–1185.

Chavan, P. V.; Dennett, K. E.; Marchand, E. A.; Gustin, M. S. Evaluation of small-scale constructed wetland for water quality and Hg transformation. *J. Hazard. Mater*. **2007**, *149(3)*, 543-547.

Danovaro, R.; Pusceddu, A. Biodiversity and ecosystem functioning in coastal lagoons: does microbial diversity play any role? *Estuar. Coast Shelf Sci*. **2007**, *75(1)*, 4-12.

Faulwetter, J. L.; Burr, M. D.; Parker, A. E.; Stein, O. R.; Camper, A. K. Influence of season and plant species on the abundance and diversity of sulfate reducing bacteria and ammonia oxidizing bacteria in constructed wetland microcosms. *Microb. Ecol*. **2013**, 65(1), 111-127.

Fisher, J.; Acreman, M. C. Wetland nutrient removal: a review of the evidence. *Hydrol. Earth Syst. Sci*. **2004**, *8(4)*, 673-685.

Jackson, E. F.; Jackson, C. R. Viruses in wetland ecosystems. *Freshwater Biol*. **2008**, *53(6)*, 1214-1227.

Jasper, J. T.; Nguyen, M. T.; Jones, Z. L.; Ismail, N. S.; Sedlak, D. L.; Sharp, J. O.; Luthy R. G; Horne, A. J.; Nelson, K. L. Unit process wetlands for removal of trace organic contaminants and pathogens from municipal wastewater effluents. *Environ. Eng. Sci*. **2013**, *30(8)*, 421-436.

Jasser, I.; Kostrzewska-Szlakowska, I.; Ejsmont-Karabin, J.; Kalinowska, K.; Weglenska, T. Autotrophic versus heterotrophic production and components of trophic chain in humic lakes: the role of microbial communities. *Pol. J. Ecol*. **2009**, *57(3)*, 423-439.

Kobayashi, T.; Ralph, T. J.; Ryder, D. S.; Hunter, S. J. Gross primary productivity of phytoplankton and planktonic respiration in inland floodplain wetlands of southeast Australia: habitat-dependent patterns and regulating processes. *Ecol. Res*. **2013**, *28(5)*, 833-843.

Li, Y.; Wang, L.; Zhang, W.; Wang, H.; Fu, X.; Le, Y. The variability of soil microbial community composition of different types of tidal wetland in Chongming Dongtan and its effect on soil microbial respiration. *Ecol. Eng*. **2011**, *37(9)*, 1276-1282.

Mitchell, C. P.; Branfireun, B. A.; Kolka, R. K. Methylmercury dynamics at the upland-peatland interface: Topographic and hydrogeochemical controls. *Water Resour. Res*. **2009**, *45(2)*, W02406.

Ramasamy, E. V.; Toms, A.; Shylesh, C. M. S.; Jayasooryan, K. K.; Mahesh, M.

Mercury fractionation in the sediments of Vembanad wetland, west coast of India. *Environ. Geochem. Health*. **2012**, *34(5)*, 575-586.

Reed, A. J.; Hicks. R. E. Microbial ecology of Lake Superior Bacteria and Archaea: An overview. *Aquat. Ecosyst. Health Manag*. **2011**, *14(4)*, 386–395.

Sims, A.; Horton, J.; Gajaraj, S.; McIntosh, S.; Miles, R. J.; Mueller, R.; Reed, R.; Hu, Z. Temporal and spatial distributions of ammonia-oxidizing archaea and bacteria and their ratio as an indicator of oligotrophic conditions in natural wetlands. *Water Res*. **2012**, *46(13)*, 4121-4129.

Suttle, C. A. Marine viruses-major players in the global ecosystem. *Nat. Rev. Microbiol*. **2007**, *5*, 801–812.

Turner, R. K.; Georgiou, S.; Brouwer, R.; Bateman, I. J.; Langford, I. J. Towards an integrated environmental assessment for wetland and catchment management. *Geogr. J*. **2003**, *169(2)*, 99-116.

Turner, R. K.; Van den Bergh, J. C. J. M.; Soderqvist, T.; Barendregt, A.; Van der Straaten, J.; Maltby, E.; Van Ierland, E. C. Ecological-economic analysis of wetlands: Scientific integration for management and policy. *Ecol. Econ*. **2000**, *35*, 7–23.

Wang, H.; Holden, J.; Spera, K.; Xu, X.; Wang, Z.; Luan, J.; Xu, X.; Zhang, Z. Phosphorus fluxes at the sediment–water interface in subtropical wetlands subjected to experimental warming: A microcosm study. *Chemosphere* **2012**, *90(6)*, 1794-1804.

Whitmire, S. L.; Hamilton, S. K. Rapid removal of nitrate and sulfate in freshwater wetland sediments. *J. Environ. Qual*. **2005**, *34*, 2062–2071.

CHAPTER 2

BACTERIOPHAGE-HOST INTERACTION IN NATURAL ECOSYSTEMS: A REVIEW OF EXPERIMENTAL AND BIOINFORMATICS ANALYSIS

2.1. Introduction

Bacteriophages (also called phages for short) are viruses that infect only bacteria as their hosts. In fact, "phage" is originated from a Greek word meaning "devour". Bacteriophages are considered as the most abundant entities in the biotic and abiotic environments with an estimated population of 10^{31} phages on the earth which outnumber coexisting bacteria by a factor of 10 (Thomas et al., 2011; Williamson et al., 2013; Engelhardt et al., 2014). For instance, there are about 1.5×10^{7} phages per gram of soil (Ashelford et al., 2003), 10^{7} phages per ml of sea water (Mann, 2005), and other living beings such as human gastrointestinal tract contain about 10^{12} phages. Since the discovery of phages by Twort (1915), there were various milestones in the last century that enliven bacteriophage study, especially in the last decade. The second discovery of phages in the early 1940s focused on *Escherichia coli* strain B as one host bacterium (Brock, 1990). Environmental microbiologist discovered phages in natural ecosystems in the 1970s as the most abundant biological entities in the aquatic global nutrients cycle by infection of a significant amount of bacterial communities and keeping the

nutrients in a microbial loop, which affects bacterial population and diversity (Fuhrman, 1999). Renaissance of antibiotic resistance bacteria (ARB) in the 1990s rediscovered the phage therapy to confront the antibiotic crisis in medical applications (Sulakvelidze et al., 2001). Furthermore, in the last few years, the human micriobiome project leads scientists to seek the potential importance of the virome for understanding the dynamics of the gut microbiota (Reyes et al., 2012).

The fundamental question that we are trying to address in this literature review chapter is "what is the role of bacteriophage in bacterial diversity, population and characteristics in a natural ecosystem?" To answer this question, we need to understand the importance of phages in natural ecosystems, understand the approaches to study the phages, and their interaction with their bacterial hosts in natural ecosystems. Therefore, we organized this review in a hierarchical manner, moving from history of phage research in the last century, through understanding of phage-host interaction in the laboratory, and complex communities in natural environments with genomics and metagenomics analysis. In the first section of this review, an overview of experimental approaches for bacteriophage abundance, diversity, and their interaction with a bacterial host in a laboratory environment and various natural ecosystems are discussed. In the second section, sequencing strategies and recent studies employing high-throughput sequencing technologies for better analysis of phage-bacteria interaction are discussed. In the last section, future research directions are recommended for new studies with more in-depth understanding of the phage-host interaction.

2.2. Bacteriophages and bacteria diversity in natural environments

One of the fundamental steps in understanding the role of phages in driving bacterial evolution and their interaction is quantifying bacteriophage and bacterial host population in various natural ecosystems. Furthermore, phage diversity and their different infection cycles are critical to understand for further studies of the phage-host interaction. Bacteria are distributed and found across most environments and niches on the planet experiencing from mild to extreme temperature, pressure, and salinity. Prior to the advent of culture-independent quantification techniques, the bacterial populations in the environment were significantly underestimated due to a very narrow range of environmental conditions in the lab. Advances in direct-count methods such as staining followed by microscopy dramatically increased the number of bacteria discovered in the environment (Hobbie et al., 1977). Environmental sequencing employing nucleic acid sequencing of 16S rRNA, i.e. the most conserved region in all bacteria, increased the resolution of environmental quantification of bacteria (Torsvik and Ovreas, 2002).

In contrast to bacterial ecology, although phages are expected to be similarly distributed as their bacterial hosts, knowledge about phage distribution in natural environments is still limited. Bacteriophages are found and abundant in various environments including fresh water (Fuhrman, 1999), marine environments (Breitbart and Rohwer, 2005), hypersaline environments (Sime-Ngando et al., 2011), soil (Salifu et al., 2013), deserts (Fancello et al., 2013), and Polar Regions (Sawstrom et al., 2008).

Similar to bacterial quantification, culture-dependent techniques in bacteriophage quantification greatly limited phage detection and significantly underestimated the phage population in the environment. Advances in using specific fluorescent staining, such as

SYBR gold followed by microscopy and flow cytometry, increased the resolution of environmental quantification of bacteriophages. In addition, culture-independent sequence identification is unlikely for phage quantification because of not sharing a conserved gene region (Rohwer and Edwards, 2002). However, with the advent of metagenomics technologies, sequence-based estimates of phage abundance revealed the more accurate extent of phage population. With the metagenomics analysis, it has been studied that the phages are ubiquitous across various natural ecosystems and can account for up to 20% turnover of the living biomass in the sea (Breitbart and Rohwer, 2005).

2.3. Bacteriophage life cycles, properties and advantages

The genetic material of bacteriophages is contained in a protein capsid, which has an icosahedral construction (Lodish et al., 2000). The capsid is connected through a collar to a tail composed of tail fibers with tips (Orlova, 2012). The tail fibers recognize receptors on the bacterial cell surface, which determine the range of a phage's host organisms. Depending on the specificity of the receptors, phages can be categorized into **monovalent phages** with the capacity to infect a specific bacterial species or strain, and **polyvalent phages** that are able to adsorb to bacterial surfaces, enabling them to infect across different bacterial species or even genera.

Bacteriophages can be categorized into two major classes according to their life cycle: virulent (lytic) phages and temperate (lysogenic) phages (Guttman et al., 2004). As shown in Figure 2.1, during a lytic cycle, virulent phage particles attach to the host bacterial surface, inject their genetic material into the host's DNA, replicate their DNA using host machinery and then assemble their structural components until lysing the host cell and releasing new phage progenies into the environment. During the lysogenic cycle,

temperate phages integrate their genome onto the host chromosome, remain as a prophage within the bacterium, and vertically transmit, by cell division, its progeny until the lytic cycle is induced (Motlagh et al., 2015).

Several life cycles are employed by phages to complete their infection process depending on the bacterial hosts. The phage life cycles can be in a range of destructive types involving bacterial cell lysis to those neutral and even beneficial to the bacterial hosts. Lytic phages infect their host by recognizing receptors on the cell surface, attaching to the receptor in a "lock and key" fashion, adsorbing to the bacterial cells, and ejecting their genetic material into the cell (Molineux and Panja, 2013). As mentioned, the adsorption process requires recognition of bacterial receptors including lipopolysaccharides, flagella, or pili on the cell surface (Lindberg, 1973; Samuel et al., 1999; Mattick, 2002). The bacterial cell machinery produces the viral proteins replicating the viral genetic materials, which are self-assembled and packaged into capsids (Aksyuk and Rossmann, 2011). After producing enough virions, the cell lyses and releases progeny viruses to start a new infection process of another susceptible host.

Temperate phages in the lysogenic cycles are capable of integrating their genetic material into the bacterial chromosome and transmit to daughter cells through bacterial reproduction. Under certain conditions, temperate phages can alter to lytic cycle that eventually lyses the cell. During coevolution of phages, prophages can lose their ability to excise from the host genome and form new virions, which is considered as crytic prophage.

In addition to the lytic and lysogenic cycles as the most common phage life cycles, pseudolysogeny is the less well-understood phage lifestyles, which has important

implications for phage-bacteria interactions in nature. In the pseudolysogeny, the phage neither integrates into the host genome nor enters the lytic cycle, which is associated with conditions of cell starvation. Since this lifestyle usually resolves with phages entering the lytic or lysogenic cycle (Los and Wegrzyn, 2011), it is not clearly understood whether considering pseudolysogeny as a life cycle or a necessary "pause" due to reduced activity of phages as a result of starvation. As nutrients might be transiently limited in the natural environment and starvation can occur frequently, phages may not decay from a population by a lack of bacterial growth and they can start their activity once cell growth resumes.

2.4. Bacteriophage and bacterial host interaction in natural environments

Beyond the distribution of phages and their bacterial hosts in the natural environments, an important parameter of studying the bacteria and phage is their ability to interact biologically. Microbial interactions in a natural ecosystem can shape biological and geochemical processes at a global scale. In particular, bacteria are predominant organisms in the microbial ecosystem and directly shape biogeochemical cycles (Karl et al., 1997). Bacterial communities can be involved in oxygen generation, carbon fixation, nitrogen fixation, ammonification, anammox process, phosphorous metabolism, and metal transformation, which results in complex biogeochemical processes in the natural ecosystem (Figure 2.2).

As many of the biogeochemical processes in the environment are emerged from bacterial communities, bacteriophages can affect the abiotic and biotic environments by altering various biogeochemical cycles (Suttle, 2007). Bacterial lysis is one of the main

ways that viruses can affect the nutrient cycles, reported to cause 50% or greater mortality in some environments (Wilhelm and Suttle, 1999). Following bacterial lysis, bacterial organic matters including nucleic acids, proteins, and lipids are released to the environments, which significantly increase the nutrient pools and the availability of trace elements. The increase in the trace elements, particularly iron, limits primary productivity in the ecosystem. Additionally, primary production can be directly limited by phages through bacterial infection up to 78% in seawater (Suttle, 1994).

Furthermore, the phage-host interaction has the potential to shape bacterial populations and communities, alter competition among bacterial species (Bohannan and Lenski, 2000a, b; Joo et al., 2006; Koskella et al., 2012), diversity of bacterial strains (Buckling and Rainey, 2002a, b; Rodriguez-Valera et al., 2009), and mediate horizontal gene transfer among bacteria (Kidambi et al., 1994; Canchaya et al., 2003). The phage-host coevolution has important effects in shaping genotypic, phenotypic, and community-level diversity (Weinbauer and Rassoulzadegan, 2004; Avrani et al., 2011). However, its studies have been hindered due to lack of understanding of the interaction networks of the phages and their hosts.

2.4.1. Dynamic changes in bacterial and viral communities abundance and diversity

Bacteriophages can have a significant impact on the bacterial abundance, especially when phages are in their lytic cycles. Phages can outnumber their bacterial hosts by a ratio of 1:10 in natural environments, which may explain a significant factor in bacterial mortality (Fuhrman and Noble, 1995). However, the oscillating cycles of bacterial and viral abundance can be considered more complex involving bacterial

resistance that can buffer the effects of phage predation (Harcombe and Bull, 2005). For instance, some phages such as *E. coli* T4 multiply 50 times faster than its *E. coli* bacterial host. With this significant difference in the host generation and infection rate, T4 phage could drive the host into extinction and then itself (Brussow, 2013). In addition, typically phages have shorter generation time and larger population size than their bacterial hosts, which would enhance the phage infection process on their hosts.

Besides the lysis process, temperate phages can also affect the bacterial population by reducing competition for nearby hosts of the same genotype. Although the bacterial cells having inducible prophage are typically lysed in the process, the rest of the bacterial population that are protected from further infection due to sharing the prophage on their genomes benefits from reduction in the competing bacterial population (Bossi et al., 2003).

In summary, phage-bacteria interaction in different environments results in dynamic changes in the phages and bacteria population. In aquaculture environments, bacteria and phages were shown to be highly dynamic at the strain level and relatively static at the community level (Rodriguez-Brito et al., 2010). On the other hand, in enriched bioreactor environments, frequent changes were found to the microbial communities with correlated variation in the phage population associated with dominant taxa (Shapiro et al., 2010). Additionally, the increase in the phages was found to correlate with reduction in the dominant bacterial taxa in mixed bacterial communities cultured in seawater (Fuhrman and Schwalbach, 2003).

2.4.2. Phage adaptation to bacterial population

Although it has been traditionally considered that the phages are highly host-specific to individual bacterial species or even strains, due to the interaction and coevolution between the phages and their hosts in the environment, it is suggested that bacterial susceptibility and resistance may not be as highly host-specific as previously supposed (Koskella and Meaden, 2013). Based on the experimental evidence from natural and laboratory systems, it is suggested that many phages may infect more than one bacterial strain or species and many bacteria can be the host to a number of different phages in the environment (Flores et al., 2011). A meta-analysis of the lytic phage infection network in nature revealed that phages with broad host range are more capable of infecting highly resistant hosts, wherein phages with narrow host range are only capable of infecting the most susceptible hosts (Flores et al., 2011). The phage specificity in the environment is supported by the locally adaptation mechanisms of phages to their bacterial hosts across space (Vos et al., 2009).

Similarly, multiple phages can infect a single bacterial genotype in the local environment. Therefore, this interaction between the phages and their host and evolution of resistance may result to changes in various phage lineages simultaneously. In addition, the horizontal gene transfer process mediated by phages (Canchaya et al., 2003; Smillie et al., 2011) and potential for recombination among phage genomes during coinfection increase the chance of phage adaptation to the bacterial populations and altering the viral host range. Furthermore, phages with greater niche breath and capable of infecting multiple susceptible bacterial hosts are expected to survive extinctions (Ogbunugafor et al., 2010). Therefore, instead of being limited to a particular bacterial host, phages can be

under selection to increase host range to enhance geographic distribution.

Bacteriophages can expand their distribution by increasing their tolerance to environmental degradation (Diaz-Munoz and Koskella, 2014). In fact, phages are generally stable under optimal conditions for long periods of time (Clokie et al., 2011) and can tolerate a range of environmental stresses greater than their hosts (Sano et al., 2004). This enhanced environmental tolerance may increase their ability to find different susceptible bacterial hosts. Interestingly, bacteriophages may adapt to environmental stress independently of their hosts' adaptation such as tolerance of phage Φ6 to heat shock at 50°C where *Pseudomonas syringae* pv. phaseolicola bacterial host for phage Φ6 cannot grow and survive at temperatures greater than 30°C (Dessau et al., 2012).

2.4.3. Phage-mediated bacterial gene transfer

Although lytic phages reproduce with little interaction between the host and viral genome and the bacterial cells' genetic machinery is merely used to reproduce the viral genetic material, in some cases, phage capsids package that are being derived from cells' nucleic acid are transported to other bacterial cells in a process called transduction. Through generalized transduction, phages can facilitate the transfer of genes among bacterial strains or genera, and therefore, the viral population can serve as a reservoir of the genes. On the other hand, via specialized transduction, specific regions of the bacterial genome are acquired for transferring into the phage genome. These targeted bacterial nucleic acids may benefit the phage by changing the host physiology and increasing its infection (Diaz-Munoz and Koskella, 2014). For instance, cyanophages carry essential genes for photosynthesis that can be used to maintain energy generation for the host cyanobacteria, and therefore, allows phages to continue reproduction even

after the host cell has terminated production of essential photosynthesis proteins (Clokie et al., 2006). Furthermore, the acquired bacterial nucleic acids may confer an advantage to the bacteria by supporting the bacterial fitness and transferring toxin-encoding genes that enhance the bacterial virulence. For instance, toxin genes encoded by phages are found in *Vibiro cholera* (Waldor and Mekalanos, 1996), *Corynebacterium diphtheriae* (Holmes and Barksdale, 1969), *Vibrio harveyi* (Munro et al., 2003), *Pseudomonas aeruginosa* (Hayashi et al., 1990), *Escherichia coli* (Newland et al., 1987), and *Shigella dysenteriae* (McDonough and Butterton, 1999) that enhance bacterial host virulence.

2.5. Bacterial response and resistance to phage infection

Bacterial fitness requires survival, reproduction, and resistance of bacteria to the phage infection, which carries substantial fitness costs to the bacterial host. These fitness costs include an increased cost of deleterious mutations (Buckling et al., 2006), decreased ability to metabolize carbon (Middelboe et al., 2009), altered competitive ability (Brockhurst et al., 2005; Lennon et al., 2007; Quance and Travisano, 2009), and increased susceptibility to other phages (Avrani et al., 2011; Marston et al., 2012). Therefore, bacterial hosts respond to the phage infection by resistance processes, which can be categorized into two main types of pre-entry resistance mechanism including adsorption reduction and postentry resistance mechanisms including restriction following the infection and abortive infections (Hyman and Abedon, 2010).

2.5.1. Phage adsorption reduction mechanism

As mentioned earlier, phages need to first attach to a specific cell-surface receptor in order to successfully infect a bacterium. To prevent phage binding, one simple way is

to lose the receptor such as in *Escherichia coli* for resistance to χ-phage (Samuel et al., 1999) or *Pseudomonas aeruginosa* (Mattick, 2002) that resistance can be readily achieved by losing or inactivating the flagella or type IV pili. Although there are a number of ways that bacteria employ to reduce the phage adsorption such as flagella loss, pilus loss, pilus modification, and hyperpiliation, since these receptors have an important function for bacterial survival, reproduction, and pathogenesis, most of the resistance mechanisms are associated with substantial fitness costs and constrained in natural environments (Bohannan and Lenski, 2000b).

An alternative resistance strategy with less fitness loss for the bacterial host is receptor modification, which results in less substantial cost than complete loss of function. For instance, an alteration to outer membrane lipopolysaccharides (LPS) in Gram-negative bacteria such as *Escherichia coli* results in bacterial resistance to T7 bacteriophage (Qimron et al., 2006).

Another strategy employed by bacteria to provide resistance to phage infection is the production of an extracellular matrix (Labrie et al., 2010). Based on spatial heterogeneity where bacteria create self-organized refuges (i.e. biofilm), within that zone, phages cannot easily infect the bacteria (Heilmann et al., 2012). On the other hand, phages survive to infect enough susceptible bacteria at the boundaries of the biofilms. Most bacteria in nature grow in the form of biofilm aggregating to itself, to other bacteria (Motlagh et al. 2013), and/or to organic and inorganic surfaces (De Oliveira et al., 2014), which provides a physical barrier for the extraneous material to penetrate. Although many bacteria are surrounded by a layer of extracellular polysaccharide (EPS) even in planktonic mode (Zaky et al., 2012), there are some phages such as T7-like (Cornelissen

et al., 2011), *Delftia tsuruhantesis* (Bhattacharjee et al., 2014), and lambdoid *Psedomonas* (Cornelissen et al., 2012) phages that contain virion-associated EPS depolymerases that can degrade the biofilm and reach the bacterial phage receptor.

2.5.2. Postentry resistance mechanisms

Following phage attachment to the bacterial host, the genetic material is injected and replication inside the bacterial cell will be started. Bacteria can also employ post-infection processes, which block on cell takeover and target and degrade phage nucleic acids as the foreign genetic material upon entry into the cell. CRISPR-Cas system (Jansen et al., 2002) and restriction enzymes (Hyman and Abedon, 2010) are two of the bacterial defense systems that may be less constrained than mechanisms targeting phage adsorption in the natural environments.

Clustered Interspaced Short Palindromic Repeats (CRISPRs) and associated genes (*CRISPR-Cas*) play an important role in targeting, neutralizing, and degrading foreign nucleic acids from phage and plasmid sources (Deveau et al., 2010). These genomic regions, typically 21-47 base-pairs (bp) in length, consisting of tandem repeated DNA sequences separated by nonrepetitive spacer sequences with approximately same length, occur in half of bacterial genomes (Godde and Bickerton, 2006). Short spacer sequences that correspond to specific sequences of the targeted nucleic acid are acquired by the CRISPR loci, which bring resistance to the target genetic material. A study conducted by Kuno et al. (2012) revealed highly diverse CRISPR sequences in the host population as an indicator of strong association and rapid evolution of the bacterial and phage populations. In a similar study, *Streptococcus thermophiles* phage system showed acquisition of at least one phage-related spacer in all of the bacterial host cells after only

one week of co-culturing (Sun et al., 2013). Additionally, special associations between CRISPR spacers and local phages were studied in natural communities across multiple sites (Held and Whitaker, 2009), which can be referred as a part of coevolution.

Although CRISPR-mediated resistance requires bacteria to specifically match the targeted phage (Barrangou et al., 2007), phages can readily evolve to escape recognition by CRISPR resistance (Levin et al., 2013; Sun et al., 2013). Anti-CRISPR counter-measures by phages has been studied and showed that phages can also carry a CRISPR-cas system to target chromosomal island of the bacterial host (Seed et al., 2013). Matching the CRISPR spacers in bacterial genomes to phage metagenome found from the same location demonstrated that phages are capable of rapidly adapting to their bacterial host (Rho et al., 2012). A similar approach investigated adaptation of *Candidatus Accumulibacter phosphatis* across two geographically separated bioreactors and showed rapid divergence of CRISPR sequences which correlated with local phage-mediated selection (Kunin et al., 2008).

The other alternative as the restriction mechanism following the phage infection is producing endonucleases by bacterial host to restrict unmethylated nucleic acid at specific sites, while protecting its own nucleic acid by methylation (Tock and Dryden, 2005). As a result, these endonucleases can degrade phage DNA and avoid them from further replication. Similar to other restriction mechanisms, phages have readily evolved various mechanisms to avoid these restriction enzymes, including loss of restriction sites, evolution of their own methylase genes, and use of unusual bases such as hydroxymethyl cytosine instead of cytosine in bacteriophage T4 (Labrie et al., 2010).

The potential impact of narrow or broad host range on the apparent competition

among bacterial hosts showed that the host resistance to the phage directly affects the dominant bacteria in any environment (Bohannan and Lenski, 2000a, 2000b). Following this model with the assumption of having a phage with two bacterial hosts, three scenarios can occur depending on the phages and their hosts; (i) the phage infects only one of the two bacterial hosts and the resistant host will become the dominant strain or species in the environment, (ii) the phage infects two bacterial hosts with different sensitivity either in terms of adsorption rate or time to reach cell lysis, and as expected, the more resistant strain will become dominant in the environment, or (iii) the bacterial hosts are able to evolve resistance to the shared phage and their relative abundance and fitness may vary depending on the cost for the resistance that each bacteria has to pay. Therefore, it is clear that phages can alter and maintain bacterial population and diversity from the range of genome level to population level and even up to microbial communities. The prevalence of these bacterial resistance mechanisms can now be productively studied using sequence-based technologies such as metagenomics and undoubtedly will shed light into phage-bacteria interaction in the environment.

In addition to bacterial resistance mechanisms to phage infection, ecological variables and environmental conditions can also mediate the phage-host interaction and alter the costs and benefits of resistance for the bacterial host, which ultimately promotes or constrains coevolutionary dynamics. Dispersal or mixing of bacterial populations can increase contact rates between the phage and host, which enhances phage transmission, allows phages to penetrate bacterial refugia and increases the coevolution rate that results in the evolution of bacteria with broader resistance and phages with extensive infectivity ranges (Brockhurst et al., 2003). In addition, rate of resources supply to bacterial hosts as

concentration of carbon substrates reduces the cost of resistance mutations and intensifies the mutation rate and increases the populations (Harrison et al., 2013). On the other hand, bacterial population also increases with resource supply, which results in elevated phage-bacteria contact rates that eventually enhance the benefits of resistance (Lopez-Pascua and Buckling, 2008). Although there were various experimental laboratory studies on the phage-bacteria interaction and coevolution (Middelboe et al., 2001; Friman and Buckling, 2013; Koskella, 2014), it is not straightforward to apply these results to make a prediction about the coevolution processes in natural microbial communities. The additional abiotic and biotic selection pressures, variation in resources, competition among species, differing microbial population sizes, and migration rate are a few of the possible features in the natural ecosystem that can dramatically alter the phage-bacteria encounter in natural environments compared to simple microcosm experiments in laboratory conditions.

2.6. Sequencing analysis to understand phage-bacteria interaction

Microbial communities including bacterial and viral population co-occur in a single habitat in natural ecosystem. The advent of cultivation-independent tools in the last few decades significantly contributed to the population and diversity studies of microbial communities. Additionally, the application of next-generation sequencing (NGS) to sequence of environmental samples, i.e. metagenomics, enhances the microbial identification resolution to large-scale insights into functional level of complex microbial communities. In this section, we have reviewed the various procedures and available pipelines for studying the genetic and metabolic networks of phage-host interaction through sequencing analysis. First, bacteria and phage population and diversity

characterization using 16S rRNA genes are discussed and then we will review the methods transforming high-throughput short sequence reads into taxonomic and functional entities to understand the interaction between the phages and their hosts.

2.6.1. 16S rRNA gene profile

The small ribosome subunit 16S gene (16S rRNA gene) can be used as a highly conserved housekeeping genetic marker between different bacterial species to study their phylogeny and taxonomy. The 16S rRNA gene contains nine hypervariable regions (V1 to V9) and 'universal' primers can be used to amplify these highly variable regions and distinguish between bacterial strains. The most significant limitation of 16S rRNA sequencing is the biases by selection of hypervariable regions as well as biases by PCR primer designs, which can result in underestimation or overestimation of particular groups of microorganisms. In addition, as the 16S rRNA gene is present in different copy numbers in bacterial genomes, the relative abundance of microorganisms will be affected (Klappenbach et al., 2000).

Bacteriophages have the lack of highly conserved genes common to all representatives and therefore, phylogenetic analysis of phages with 16S rRNA sequencing is a challenging process. When there is background knowledge about the viral families present in the sample or whether the presence of a particular phage-encoded gene is known, PCR can be the possible technique (Casas et al., 2006). However, the choice of gene to be amplified and the consequent design of primers are the biggest challenge in phage PCR and sequencing. For phage diversity, T4-like viruses in the family of *Myoviridae* are among the most understood bacteriophage genus, which have been used in previous studies of cyanophages (Fuller et al., 1998; Marston and Sallee,

2003; Short and Suttle, 2005) and marine viral diversity (Filee et al., 2005). The primers for these phage PCR have been designed for the major capsid proteins (gp23 and gp20), and DNA polymerase encoding gene (gp43). The interaction between phage and bacteria might results in viral diversity inaccuracy when screening for phages' encoded genes acquired from their bacterial hosts. For instance, cyanophages that acquired photosynthesis genes may be equally amplified with the same primers from *Podoviridae* or *Myoviridae* families (Sullivan et al., 2006).

2.6.2. Bioinformatics analysis for bacterial and viral metagenomics

The sequencing of 16S rRNA gene mainly identifies the abundance and diversity of bacteria and phages in environmental samples. However, the metagenomic analysis additionally can identify the gene content, potential of functional proteins encoded in the genomes, novel enzymatic functions, host-pathogen interaction, and even novel therapeutic strategies in human diseases. Advent and advance in high-throughput sequencing technologies in the last decade have significantly improved our understanding of the microbial communities in the natural ecosystems. In the following sections, we reviewed various studies and approaches using high-throughput technologies to better analyze and understand the phage-bacterial interaction in natural environments.

One of the advantages of high-throughput sequencing of phage genomes and metagenomes is generating more sequence data and expanding the phage sequence libraries that can contribute to a better probing of phageom with PCR-based approaches. The initial step for metagenomics sequencing is the isolation of the total nucleic acid from the environmental samples. Prior to viral DNA extraction, the sediment samples containing phages are suspended in an osmotic neutral buffer such as SMG buffer

followed by vigorous shaking to resuspend the viral particles. In order to separate the viral particles from the microbial cells, ultracentrifugation with cesium chloride (CsCl) density gradient can be performed to purify the extracted phage sample. Furthermore, in order to remove the residual nonencapsulated free nucleic acids, DNase and RNase treatment are carried out on the purified sample.

In some environmental samples, bacterial population might not be abundant enough to extract sufficient nucleic acids. In addition, although phages outnumber bacterial cells in a ratio of 10:1 in most environments, viral DNA represents only 0.1% of the total DNA in a microbial community (Qin et al., 2010). Therefore, in many environmental samples, the amplification of total DNA is typically necessary in order to increase the nucleic acid yield to a sufficient quantity for sequencing. For this purpose, a range of amplification methods for bacterial and viral samples have been developed such as random amplified shotgun library (RASL) (Rohwer et al., 2001), linker-amplified shotgun library (LASL) (Breitbart et al., 2003), and multiple displacement amplification (MDA) (Hutchison et al., 2005). However, because of random amplification of nucleic acid in these processes, quantitative biases in sequencing may be generated and association of phage taxa to bacterial taxa will be challenging to establish. Therefore, to minimize any biases and circumvent production of artifacts such as chimeras, recent studies avoid application of any amplification process for augmentation of nucleic acid concentration.

Depending on the selected sequencing platform, a DNA library will be constructed for sequence analysis by fragmentation of the total DNA into smaller pieces of DNA followed by repairing the 5' and/or 3' ends of the DNA library. Ligation of the

adapters containing sequences for hybridization to the flow cell followed by library cleanup, amplification, and quantification are the last steps before high-throughput sequencing.

Following high-throughput sequencing, the sequences obtained by shotgun analysis are subjected to data filtering based on several quality criteria including read length, base quality, ambiguous base calls, and removing the adapter sequences. In case of pair-end sequencing when using Illumina sequencing platform, interleaving of the reads is also required that can be separately or simultaneously performed in the contig assembly process.

The quality controlled reads will be used for *de novo* assembly to generate full-length sequences as contigs and/or scaffold. Although mapping against known genomes as existing backbone sequence can recover complete genomes, usually because of the presence of multiple closely related strains, recovering complete genomes will be a highly challenging metagenomic approach. Metagenomic assemblers generally use two major approaches for sequence assembly. Graph-based reconstruction assembly using overlap-layout-consensus (OLC) to assemble longer sequences in software such as MIRA and Celera is one approach in sequence assembly. The second approach is de Brujin graph to assemble shorter sequences in software such as Euler, Velvet, ABySS, and SOAP *de novo*, which has shown more success in shorter sequences produced by Illumina and Ion Torrent platforms.

The assembly algorithm may results in erroneous chimeric contigs by assembly of two different sequences from different organisms or species. This issue is particularly relevant to viral metagenomics where the sample may contain a multitude of related viral

sequences. In order to minimize this problem and validate the assembly process, it is suggested that hybrid assembly by using several assembly algorithms needs to be employed.

The taxonomic classification and binning can be accomplished prior to or after the sequence assembly using similarity-based and non-similarity-based approaches. For similarity-based taxonomic classification, the assembled contigs should be compared against gene markers using BLAST (Segata et al., 2012), USEARCH (Edgar, 2010), or HMMs (Finn et al., 2011) to annotate and quantify each metagenomics homologue. Depending on the targeted study, the sequence identity of the best match determines the nearest phylogenetic origin of the reads. In case of using BLAST search, BLASTn algorithm may overestimate the matched hit of the nucleotide sequences and therefore, it is more preferable to search against the protein database using BLASTx or tBLASTx that translate query and reference nucleotide sequences in all six frames and then compare them to each other. As conventional BLAST-based search tools particularly BLASTx and tBLASTx are extremely time consuming when working with large metagenomic datasets, parallel implementation of BLAST can achieve super-linear speed-up up to 100 times by distributing the work across the cluster's nodes, executing searches on multiple non-shared memory processors and reassembling the results in software such as open-source parallel BLAST (mpiBLAST) (Darling et al., 2003) or commercial Paracel BLAST software, which is almost five times faster than mpiBLAST (Boysen and Riefell, 2004). In order to classify the matched sequences after the alignment process, various software including MEGAN, GAAS, GRAMMy, and GASiC can be used. Classification of viruses is performed according to the National Center for Biotechnology Information

(NCBI) viral classification system, as proposed originally by the International Committee on Taxonomy of Viruses.

Taxonomic classification using nonsimilarity approaches, i.e. composition-based method explore composition of genome such as GC content, codon usage, or k-mers usage. This taxonomic classification approach is particularly beneficial when the sequences do not have any homologs such as novel phages or they are highly divergent from sequences in public databases. Although composition-based methods are computationally faster than similarity-based methods, they have lower accuracy and are highly dependent on sequence length. Composition-based taxonomic classification can be performed using software with assignment dependent such as PhyloPythia and Phymm; and assignment independent such as Metacluster, TETRA, variants of SOMs, CompostBin, and AbundanceBin. Most of the mentioned taxonomic profiling tools are designed for analyzing bacterial metagenomic and they could have issues in analyzing viral metagenomic datasets for reasonable profile and estimation of viral sequences abundance (Klingenberg et al., 2013). Finally, it should be mentioned that besides the similarity-based and composition-based methods, some taxonomic classification softwares are hybrid using both approaches such as SPHINX and PhymmBL.

Microbial community structure and their difference between various metagenomic samples can identify the influences of microbial communities' patterns and phage-host interaction. For bacterial and viral metagenome diversity and community structure estimation, average contig spectra is generated by tools such as Circonspect, while average genome length is calculated by software such as GAAS, and ultimately, biodiversity can be estimated using these parameters by tools like PHACCS.

Microbial community diversity is calculated based on the assumption that more abundant organisms have longer and higher coverage contigs. Therefore, microbial richness (alpha diversity) is estimated by the number of different genotypes in the sample, and microbial evenness (gamma diversity) is estimated by the relative abundances and distribution of genotypes among different metagenomic samples.

The functional diversity of the bacterial and viral samples can be understood by identifying the protein coding genes and annotating them against various protein databases. Following the sequence assembly, the open reading frames (ORFs) will be predicted from the contigs using tools such as Glimmer-MG, MetaGeneMarks, and Prodigal. The predicted ORFs can be compared against protein databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG), protein family annotations (PFAM), gene ontologies (GO), clusters of orthologous groups (COG), and phage SEED.

MetaGeneAnnotator (MGA) precisely predicts all kinds of prokaryotic genes by integrating statistical models for prophage genes as well as bacterial and archaeal genes. MGA also uses a self-training model from input sequences based on GC content, which detects both typical and atypical genes such as horizontally transferred and prophage genes in prokaryotic genomes. The other method of gene prediction is by using FragGeneScan, which incorporates in Hidden Markov Model (HMM) implementing codon usage bias and start/stop codon patterns.

Besides the step-by-step procedures that were mentioned in this section, there are some automated platforms with integrated suites that have been designed to analyze bacterial and viral metagenomics sequences. Platforms such as MG-RAST, IMG/M, JCVI Metagenomics Reports (METAREP) for bacterial metagenomics and METAVIR

for viral metagenomics are among automated platforms, which are being widely used.

2.7. Future research directions

As mentioned earlier, there are multiple phages that interact and infect a single bacterial host and understanding the interaction network between these various phages capable of infecting a single bacterium is important. Phages can mediate interaction with other viruses through superinfection exclusion (Turner et al., 1999). Also, satellite phages are defective in some functions of viral cycles and their successful lysis depends on a "helper virus" to provide these aspects to reproduce (Lindqvist et al., 1997). In addition, phage-phage interaction can affect the phage-bacteria interaction as well. For instance, two distinct prophage elements can combine to produce a "composite phage" which is capable of entering lytic cycle (Duerkop et al., 2012). Some studies have been performed in a laboratory environment for understanding the interaction between multiple phages in infecting a single bacterial strain (Korona and Levin, 1993). However, although viral coinfection and phage-phage interactions occur frequently in nature (Dapalma et al., 2010), the studies related to phage-phage interaction in natural ecosystems are very limited.

The bacterial and viral metatranscriptomic can analyze the gene regulation and expression as well as dynamics of a transcriptionally active microbial community. Therefore, the application of metagenomic analysis combined with metatranscriptomic can reveal significant important correlation between the phage-host interactions. Studies are necessary to understand how bacterial metabolic activity can affect bacterial and viral fitness and how external environmental factors are implicated in their expression. Although metagenomic analysis reveals a considerable amount of information regarding

the bacterial and viral abundance, diversity, and their interaction, it could also overestimate or underestimate the functional relevance of the encoded genes if analyzed without metatranscriptomic study (Franzosa et al., 2014). It is suggested that combining phage-host metagenomic and metatranscriptomic analyses allow a deep understanding on microbial interactions with providing useful insights about the microorganisms that have relevant functions and simultaneously active genes and pathways. In addition, identification of expressed genes from prophages under different environmental conditions can provide better understanding of the interaction between prophages and bacterial host. Study of the phage gene expression in various natural ecosystems can indicate the role of encoded prophage genes in microbial physiology and determine the dynamics between active phages and their bacterial hosts.

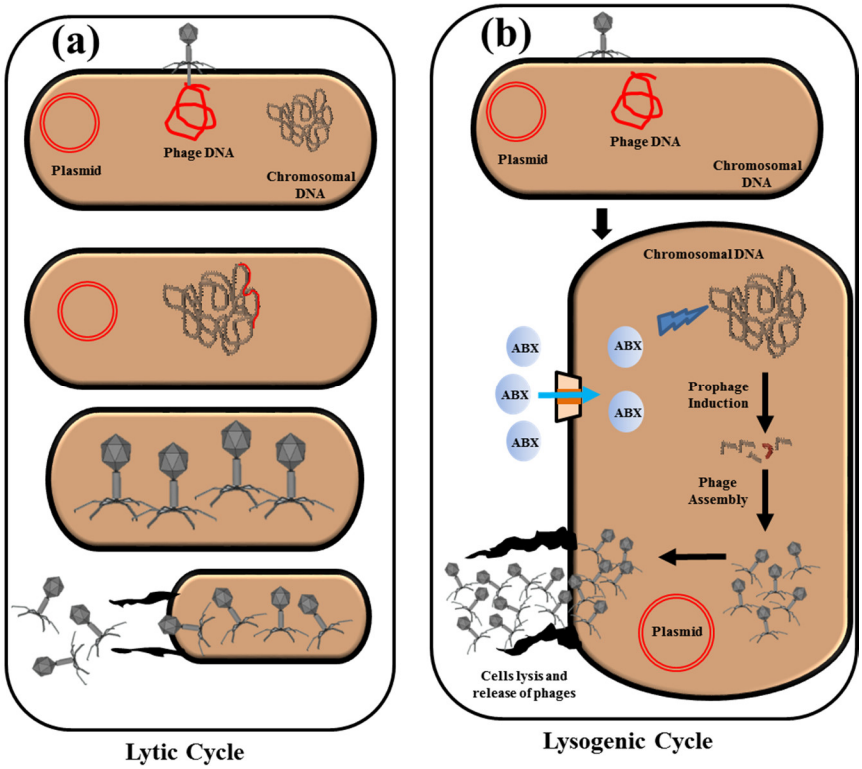

Figure 2.1. Bacteriophage life cycles (a) lytic wherein the immediate infection causes cell lysis (b) lysogenic wherein the temperate phage recombines with the host DNA forming a dormant prophage

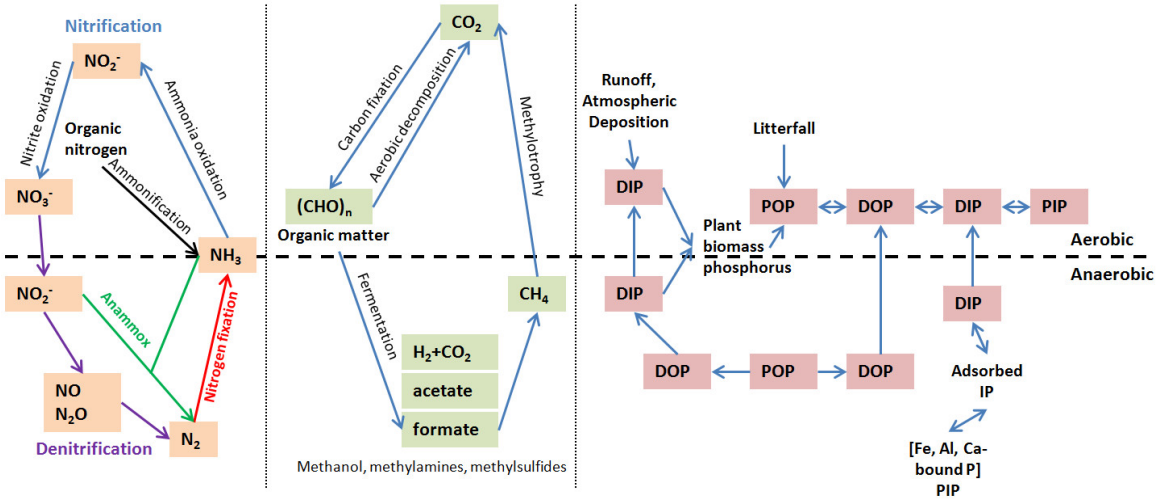

Figure 2.2. Complex biogeochemical processes of nutrients (nitrogen, carbon, and phosphorus) in aerobic and anaerobic environments of natural ecosystems

2.8. References

Aksyuk, A. A.; Rossmann, M. G. Bacteriophage assembly. *Viruses* **2011**, *3(3)*, 172-203.

Ashelford, K. E.; Day, M. J.; Fry, J. C. Elevated abundance of bacteriophage infecting bacteria in soil. *Appl. Environ. Microbiol*. **2003**, *69*, 285-289.

Avrani, S.; Wurtzel, O.; Sharon, I.; Sorek, R.; Lindell, D. Genomic island variability facilitates Prochlorococcus-virus coexistence. *Nature* **2011**, *474*, 604-608.

Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D. A.; Horvath, P. CRISPR provides acquired resistance against viruses in Prokaryotes. *Science* **2007**, *315*, 1709-1712.

Bhattacharjee, A. S.; Choi, J.; Motlagh, A. M.; Mukherji, S. T.; Goel, R. Bacteriophage therapy for membrane biofouling in membrane bioreactors and antibiotic-resistant bacterial biofilms. *Biotechnol. Bioeng*. **2015**, *112(8)*, 1644-1654.

Bohannan, B. J. M.; Lenski, R. E. The relative importance of competition and predation varies with productivity in a model community. *Am. Nat*. **2000a**, *156*, 329-340.

Bohannan, B. J. M.; Lenski, R. E. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecol. Lett*. **2000b**, *3*, 362-377.

Bossi, L.; Fuentes, J. A.; Mora, G.; Figueroa-Bossi, N. Prophage contribution to bacterial population dynamics. *J. Bacteriol*. **2003**, *185(21)*, 6467-6471.

Boysen, C.; Rieffel, M. A. Enhancing BLAST Performance by using paracel filtering package. Paracel Technology. **2004**.

Breitbart, M.; Rohwer, F. Here a virus, there a virus, everywhere the same virus? *Trends Microbiol*. **2005**, *13*, 278-284.

Breitbart, M.; Hewson, I.; Felts, B.; Mahaffy, J. M.; Nulton, J.; Salamon, P.; Rohwer, F. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol*. **2003**, *185(20)*, 6220-6223.

Brock, T. D. The Emergence of Bacterial Genetics. New York, USA: CSHL Press. **1990**.

Brockhurst, M. A.; Buckling, A.; Rainey, P. B. The effect of a bacteriophage on diversification of the opportunistic bacterial pathogen, Pseudomonas aeruginosa. *Proc. R. Soc. Lond. B Biol. Sci*. **2005**, *272*, 1385–1391.

Brockhurst, M. A.; Morgan, A. D.; Rainey, P. B.; Buckling, A. Population mixing accelerates coevolution. *Ecol. Lett*. **2003**, *6*, 975–979.

Brussow, H. Bacteriophage–host interaction: from splendid isolation into a messy reality. *Curr. Opin. Microbiol*. **2013**, *16(4)*, 500-506.

Buckling, A.; Rainey, P. B. The role of parasites in sympatric and allopatric host diversification. *Nature* **2002a**, *420*, 496-499.

Buckling, A.; Rainey, P. B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. Lond. B Biol. Sci*. **2002b**, *269*, 931-936.

Buckling, A.; Wei, Y.; Massey, R. C.; Brockhurst, M. A.; Hochberg, M. E. Antagonistic coevolution with parasites increases the cost of host deleterious mutations. *Proc. R. Soc. Lond. B Biol. Sci.* **2006**, *273(1582)*, 45-49.

Canchaya, C.; Fournous, G.; Chibani-Chennoufi, S.; Dillmann, M. L.; Brussow, H. Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol*. **2003**, *6*, 417-424.

Casas, V.; Miyake, J.; Balsley, H.; Roark, J.; Telles, S.; Leeds, S.; Zurita, I.; Breitbart, M.; Bartlett, D.; Azam, F.; Rohwer, F. Widespread occurrence of phage-encoded exotoxin genes in terrestrial and aquatic environments in Southern California. *FEMS Microbiol. Lett*. **2006**, *261(1)*, 141-149.

Clokie, M. R. J.; Millard, A. D.; Letarov, A. V.; Heaphy, S. Phages in nature. *Bacteriophage* **2011**, *1*, 31–45.

Clokie, M. R.; Shan, J.; Bailey, S.; Jia, Y.; Krisch, H. M.; West, S.; Mann, N. H. Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. *Environ. Microbiol*. **2006**, *8(5)*, 827-835.

Cornelissen, A.; Ceyssens, P. J.; Krylov, V. N.; Noben, J. P.; Volckaert, G.; Lavigne, R. Identification of EPS-degrading activity within the tail spikes of the novel Pseudomonas putida phage AF. *Virol. J*. **2012**, *434(2)*, 251-256.

Cornelissen, A.; Ceyssens, P. J.; T'syen, J.; Van Praet, H.; Noben, J. P.; Shaburova, O. V.; Krylov, V. N.; Volckaert, G.; Lavigne, R. The T7-related Pseudomonas putida phage φ15 displays virion-associated biofilm degradation properties. *PLoS One* **2011**, *6(4)*, e18597.

Darling, A.; Carey, L.; Feng, W. C. The design, implementation, and evaluation of mpiBLAST. *Proceedings of ClusterWorld* **2003**, *2003*, 13-15.

De Oliveira, L. D.; Motlagh, A. M.; Goel, R.; de Souza Missagia, B.; de Abreu Filho, B. A.; Lautenschlager, S. R. The use of bottle caps as submerged aerated filter medium. *CEP*. **2014**, *87020*, 900.

Dessau, M.; Goldhill, D.; McBride, R. L.; Turner, P. E.; Modis, Y. Selective pressure causes an RNA virus to trade reproductive fitness for increased structural and thermal

stability of a viral enzyme. *PLoS Genet*. **2012**, *8(11)*, e1003102.

Deveau, H.; Garneau, J. E.; Moineau, S. CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu. Rev. Microbiol*. **2010**, *64*, 475-493

Diaz-Munoz, S. L.; Koskella, B. Bacteria-phage interactions in natural environments. *Adv Appl. Microbiol*. **2014**, *89*, 135-83.

Duerkop, B. A.; Clements, C. V.; Rollins, D.; Rodrigues, J. L.; Hooper, L. V. A composite bacteriophage alters colonization by an intestinal commensal bacterium. *Proc. Natl. Acad. Sci*. **2012**, *109(43)*, 17621-17626.

Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26(19)*, 2460-2461.

Engelhardt, T.; Kallmeyer, J.; Cypionka, H.; Engelen, B. High virus-to-cell ratios indicate ongoing production of viruses in deep subsurface sediments. *ISME J*. **2014**, *8(7)*, 1503-1509.

Fancello, L.; Trape, S.; Robert, C.; Boyer, M.; Popgeorgiev, N.; Raoult, D.; Desnues, C. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J*. **2012**, *7(2)*, 359-369.

Filee, J.; Tetart, F.; Suttle, C. A.; Krisch, H. M. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102(35)*, 12471-12476.

Finn, R. D.; Clements, J.; Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*. **2011**, gkr367.

Flores, C. O.; Meyer, J. R.; Valverde, S.; Farr, L.; Weitz, J. S. (2011) Statistical structure of host–phage interactions. *Proc. Natl. Acad. Sci. U.S.A*. **2011**, *108*, E288–E297.

Franzosa, E. A.; Morgan, X. C.; Segata, N.; Waldron, L.; Reyes, J.; Earl, A. M.; Giannoukos, G.; Boylan, M. R.; Ciulla, D.; Gevers, D.; Izard, J.; Garrett, W. S.; Chan, A. T. Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci. U.S.A*. **2014**, *111(22)*, E2329-38.

Friman, V. P.; Buckling, A. Effects of predation on real-time host–parasite coevolutionary dynamics. *Ecol. Lett*. **2013**, *16(1)*, 39-46.

Fuhrman, J. A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541-548.

Fuhrman, J. A.; Noble, R. T. Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnol. Oceanogr*. **1995**, *40(7)*, 1236-1242.

Fuhrman, J. A.; Schwalbachm, M. Viral influence on aquatic bacterial communities. *Biol. Bull*. **2003**, *204*, 192-195

Fuller, N. J.; Wilson, W. H.; Joint, I. R.; Mann, N. H. Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Appl. Environ. Microbiol*. **1998**, *64(6)*, 2051-2060.

Godde, J. S.; Bickerton, A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol*. **2006**, *62(6)*, 718-729.

Guttman, B.; Raya, R.; Kutter, E. 3 Basic Phage Biology. *Bacteriophages: Biology and Applications* **2004**, *29*.

Harcombe, W. R.; Bull, J. J. Impact of phages on two-species bacterial communities. *Appl. Environ. Microbiol*. **2005**, *71(9)*, 5254-5259.

Harrison, E.; Laine, A. L.; Hietala, M.; Brockhurst, M. A. Rapidly fluctuating environments constrain coevolutionary arms races by impeding selective sweeps. *Proc. R. Soc. Lond. B Biol. Sci*. **2013**, *280*, 20130937.

Hayashi, T.; Baba, T.; Matsumoto, H.; Terawaki, Y. Phage-conversion of cytotoxin production in Pseudomonas aeruginosa. *Mol. Microbiol*. **1990**, *4(10)*, 1703-1709.

Heilmann, S.; Sneppen, K.; Krishna, S. Coexistence of phage and bacteria on the boundary of self-organized refuges. *Proc. Natl. Acad. Sci. U.S.A*. **2012**, *109*, 12828-12833.

Held, N. L.; Whitaker, R. J. Viral biogeography revealed by signatures in Sulfolobus islandicus genomes. *Environ. Microbiol*. **2009**, *11*, 457-466

Hobbie, J. E.; Daley, R. J.; Jasper, S. Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl. Environ. Microbiol*. **1977**, *33(5)*, 1225-1228.

Holmes, R. K.; Barksdale, L. Genetic Analysis of tox+ and tox Bacteriophages of Corynebacterium diphtheriae. *J. Virol*. **1969**, *3(6)*, 586-598.

Hutchison, C. A.; Smith, H. O.; Pfannkoch, C.; Venter, J. C. Cell-free cloning using φ29 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102(48)*, 17332-17336.

Hyman, P.; Abedon, S. T. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol*. **2010**, *70*, 217-248.

Jansen, R.; Embden, J.; Gaastra, W.; Schouls, L. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol*. **2002**, *43(6)*, 1565-1575.

Joo, J.; Gunny, M.; Cases, M.; Hudson, P.; Albert, R.; Harvill, E. Bacteriophage-mediated competition in Bordetella bacteria. *Proc. R. Soc. Lond. B Biol. Sci*. **2006**, *273*, 1843–1848.

Karl, D.; Letelier, R.; Tupas, L.; Dore, J.; Christian, J.; Hebel, D. The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* **1997**, *388(6642)*, 533–538.

Kidambi, S. P.; Ripp, S.; Miller, R. V. Evidence for phage-mediated gene transfer among Pseudomonas aeruginosa strains on the phylloplane. *Appl. Environ. Microbiol*. **1994**, *60*, 496–500.

Klappenbach, J. A.; Dunbar, J. M.; Schmidt, T. M. rRNA operon copy number reflects ecological strategies of bacteria. *Appl. Environ. Microbiol*. **2000**, *66(4)*, 1328-1333.

Klingenberg, H.; Abhauer, K. P.; Lingner, T.; Meinicke, P. Protein signature-based estimation of metagenomic abundances including all domains of life and viruses. *Bioinformatics* **2013**, btt077.

Korona, R.; Levin, B. R. Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution* **1993**, 556-575.

Koskella, B.; Lin, D. M.; Buckling, A.; Thompson, J. N. The costs of evolving resistance in heterogeneous parasite environments. *Proc. R. Soc. Lond. B Biol. Sci*. **2012**, *279*, 1896–1903.

Koskella, B. Bacteria-Phage Interactions across Time and Space: Merging Local Adaptation and Time-Shift Experiments to Understand Phage Evolution. *Amer. Nat*. **2014**, *184(S1)*, S9-S21.

Koskella, B.; Meaden, S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* **2013**, *5(3)*, 806-823.

Kunin, V.; He, S.; Warnecke, F.; Peterson, S. B.; Martin, H. G.; Haynes, M.; Ivanova, N.; Blackall, L. L.; Breitbart, M.; Rohwer, F.; McMahon, K. D. A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res*. **2008**, *18(2)*, 293-297.

Kuno, S.; Yoshida, T.; Kaneko, T.; Sako, Y. Intricate interactions between the bloom-forming cyanobacterium Microcystis aeruginosa and foreign genetic elements, revealed by diversified clustered regularly interspaced short palindromic repeat (CRISPR) signatures. *Appl. Environ. Microbiol*. **2012**, *78*, 5353-5360.

Labrie, S. J.; Samson, J. E.; Moineau, S. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol*. **2010**, *8*, 317-327

Lennon, J. T.; Khatana, S. A. M.; Marston, M. F.; Martiny, J. B. Is there a cost of virus resistance in marine cyanobacteria? *ISME J*. **2007**, *1*, 300-312.

Levin, B. R.; Moineau, S.; Bushman, M.; Barrangou, R. The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. *PLoS Genet*. **2013**, *9*, e1003312.

Lin, T. Y.; Lo, Y. H.; Tseng, P. W.; Chang, S. F.; Lin, Y. T.; Chen, T. S. A T3 and T7 recombinant phage acquires efficient adsorption and a broader host range. *PLoS One* **2012**, *7*, e30954.

Lindberg, A. A. Bacteriophage receptors. *Annu. Rev. Microbiol*. **1973**, *27(1)*, 205-241.

Lindqvist, B. H.; Deho, G.; Calendar, R. Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiol. Mol. Biol. Rev*. **1993**, *57(3)*, 683-702.

Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Viruses: Structure, Function, and Uses. **2000**.

Lopez-Pascua, L. D. C.; Buckling, A. Increasing productivity accelerates host-parasite coevolution. *J. Evol. Biol*. **2008**, *21*, 853–860.

Los, M.; Wegrzyn, G. Pseudolysogeny. *Adv. Virus Res*. **2011**, *82*, 339-349.

Mann, N. H. The third age of phage. *PLoS Biol*. **2005**, *3(5)*.

Marston, M. F.; Pierciey, F. J.; Shepard, A.; Gearin, G.; Qi, J.; Yandava, C.; Schuster, S. C.; Henn, M. R.; Martiny, J. B. H. Rapid diversification of coevolving marine Synechococcus and a virus. *Proc. Natl. Acad. Sci. U.S.A*. **2012**, *109*, 4544-4549.

Marston, M. F.; Sallee, J. L. Genetic diversity and temporal variation in the cyanophage community infecting marine Synechococcus species in Rhode Island's coastal waters. *Appl. Environ. Microbiol*. **2003**, *69(8)*, 4639-4647.

Mattick, J. S. Type IV pili and twitching motility. *Annu. Rev. Microbiol*. **2002**, *56(1)*, 289-314.

McDonough, M. A.; Butterton, J. R. Spontaneous tandem amplification and deletion of the Shiga toxin operon in Shigella dysenteriae 1. *Mol. Microbiol*. **1999**, *34(5)*, 1058-1069.

Middelboe, M.; Holmfeldt, K.; Riemann, L.; Nybroe, O.; Haaber, J. Bacteriophages drive strain diversification in a marine Flavobacterium: implications for phage resistance and physiological properties. *Environ. Microbiol*. **2009**, *11*, 1971-1982.

Middelboe, M.; Hagstrom, A.; Blackburn, N.; Sinn, B.; Fischer, U.; Borch, N. H.; J.

Pinhassi, J.; Simu, K.; Lorenz, M.G. Effects of bacteriophages on the population dynamics of four strains of pelagic marine bacteria. *Microb. Ecol*. **2001**, *42(3)*, 395-406.

Molineux, I. J.; Panja, D. Popping the cork: mechanisms of phage genome ejection. *Nature Rev. Microbiol*. **2013**, *11(3)*, 194-204.

Motlagh, A. M.; Bhattacharjee, A. S.; Goel, R. Microbiological study of bacteriophage induction in the presence of chemical stress factors in enhanced biological phosphorus removal (EBPR). *Water Res*. **2015**, *81*, 1-14.

Motlagh, A. M.; Pant, S.; Gruden, C. The impact of cell metabolic activity on biofilm formation and flux decline during cross-flow filtration of ultrafiltration membranes. *Desalination* **2013**, *316*, 85-90.

Munro, J.; Oakey, J.; Bromage, E.; Owens, L. Experimental bacteriophage-mediated virulence in strains of Vibrio harveyi. *Dis. Aquat. Organ*. **2003**, *54*, 187-194.

Newland, J. W.; Strockbine, N. A.; Neill, R. J. Cloning of genes for production of Escherichia coli Shiga-like toxin type II. *Infect. Immun*. **1987**, *55(11)*, 2675-2680.

Ogbunugafor, C. B.; Pease, J. B.; Turner, P. E. On the possible role of robustness in the evolution of infectious diseases. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2010**, *20(2)*, 026108.

Orlova, E. V. Bacteriophages and their structural organisation (pp. 3-30). InTech. **2012**.

Qimron, U.; Marintcheva, B.; Tabor, S.; Richardson, C. C. Genomewide screen of E. coli genes affecting growth of T7 bacteriophage. *Proc. Natl. Acad. Sci. U.S.A*. **2006**, *103*, 19039-19044.

Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K. S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; Mende, D. R. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464(7285)*, 59-65.

Quance, M. A.; Travisano, M. Effects of temperature on the fitness cost of resistance to bacteriophage T4 in Escherichia coli. *Evolution* **2009**, *63*, 1406–1416.

Reyes, A.; Semenkovich, N. P.; Whiteson, K.; Rohwer, F.; Gordon, J. I. Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature Rev. Microbiol*. **2012**, *10*, 607-617.

Rho, M.; Wu, Y. W.; Tang, H.; Doak, T. G.; Ye, Y. Diverse CRISPRs evolving in human microbiomes. *PLoS Genet*. **2012**, *8*, e1002441.

Rodriguez-Brito, B.; Li, L.; Wegley, L.; Furlan, M.; Angly, F.; Breitbart, M.; Buchanan, J.; Desnues, C.; Dinsdale, E.; Edwards, R.; Felts, B. Viral and microbial community

dynamics in four aquatic environments. *ISME J*. **2010**, *4*, 739–751.

Rodriguez-Valera, F.; Martin-Cuadrado, A. B.; Rodriguez-Brito, B.; Pasic, L.; Thingstad, T. F.; Rohwer, F.; Mira, A. Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol*. **2009**, *7*, 828-836.

Rohwer, F.; Edwards, R. The Phage Proteomic Tree: a genome-based taxonomy for phage. *J. Bacteriol*. **2002**, *184(16)*, 4529-4535.

Rohwer, F.; Seguritan, V.; Choi, D. H.; Segall, A. M.; Azam, F. Production of shotgun libraries using random amplification. *Biotechniques* **2001**, *31(1)*, 108-119.

Salifu, S. P.; Casey, S. A.; Foley, S. Isolation and characterization of soilborne virulent bacteriophages infecting the pathogen Rhodococcus equi. *J. Appl. Microbiol*. **2013**, *114(6)*, 1625-1633.

Samuel, A. D.; Pitta, T. P.; Ryu, W. S.; Danese, P. N.; Leung, E. C.; Berg, H. C. Flagellar determinants of bacterial sensitivity to χ-phage. *Proc. Natl. Acad. Sci*. **1999**, *96(17)*, 9863-9866.

Sano, E.; Carlson, S.; Wegley, L.; Rohwer, F. Movement of viruses between biomes. *Appl. Environ. Microbiol*. **2004**, *70*, 5842–5846.

Sawstrom, C.; Lisle, J.; Anesio, A. M.; Priscu, J. C.; Laybourn-Parry, J. Bacteriophage in polar inland waters. *Extremophiles* **2008**, *12(2)*, 167-175.

Seed, K. D.; Lazinski, D. W.; Calderwood, S. B.; Camilli, A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* **2013**, *494*, 489-491.

Segata, N.; Waldron, L.; Ballarini, A.; Narasimhan, V.; Jousson, O.; Huttenhower, C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **2012**, *9(8)*, 811-814.

Shapiro, O.H.; Kushmaro, A.; Brenner, A. Bacteriophage predation regulates microbial abundance and diversity in a fullscale bioreactor treating industrial wastewater. *ISME J*. **2010**, *4*, 327–336.

Short, C. M.; Suttle, C. A. Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Appl. Environ. Microbiol*. **2005**, *71(1)*, 480-486.

Sime-Ngando, T.; Lucas, S.; Robin, A.; Tucker, K. P.; Colombet, J.; Bettarel, Y.; Desmond, E.; Gribaldo, S.; Forterre, P.; Breitbart, M.; Prangishvili, D. Diversity of virus–host systems in hypersaline Lake Retba, Senegal. *Environ. Microbiol*. **2011**, *13(8)*, 1956-1972.

Smillie, C. S.; Smith, M. B.; Friedman, J.; Cordero, O. X.; David, L. A.; Alm, E. J. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **2011**, *480*, 241–244.

Sulakvelidze, A.; Alavidze, Z.; Morris, J. G. Bacteriophage therapy. *Antimicrob. Agents Chemother*. **2001**, *45(3)*, 649-659.

Sullivan, M. B.; Lindell, D.; Lee, J. A.; Thompson, L. R.; Bielawski, J. P.; Chisholm, S. W. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol*. **2006**, *4(8)*, e234.

Sun, C. L.; Barrangou, R.; Thomas, B. C.; Horvath, P.; Fremaux, C.; Banfield, J. F. Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol*. **2013**, *15*, 463-470.

Suttle, C. A. Marine viruses - major players in the global ecosystem. *Nature Rev. Microbiol*. **2007**, *5*, 801–812.

Suttle, C. A. The significance of viruses to mortality in aquatic microbial communities. *Microb. Ecol*. **1994**, *28(2)*, 237-243.

Thomas, R.; Berdjeb, L.; Sime-Ngando, T.; Jacquet, S. Viral abundance, production, decay rates and life strategies (lysogeny versus lysis) in Lake Bourget (France). *Environ. Microbiol*. **2011**, *13*, 616–630.

Tock, M. R.; Dryden, D. T. F. The biology of restriction and anti-restriction. *Curr. Opin. Microbiol*. **2005**, *8*, 466-472.

Torsvik, V.; Ovreas, L. Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol*. **2002**, *5(3)*, 240-245.

Turner, P. E.; Burch, C. L.; Hanley, K. A.; Chao, L. Hybrid frequencies confirm limit to coinfection in the RNA bacteriophage φ6. *J. Virol*. **1999**, *73(3)*, 2420-2424.

Twort, F. W. An investigation on the nature of ultra-microscopic viruses. *Lancet* **1915**, *2*, 1241-1243.

Vos, M.; Birkett, P. J.; Birch, E.; Griffiths, R. I.; Buckling, A. Local adaptation of bacteriophages to their bacterial hosts in soil. *Science* **2009**, *325(5942)*, 833-833.

Waldor, M. K.; Mekalanos, J. J. Lysogenic conversion by a filamentous phage encoding Cholera toxin. *Science* **1996**, *272*, 1910–1914.

Weinbauer, M. G.; Rassoulzadegan, F. Are viruses driving microbial diversification and diversity? *Environ. Microbiol*. **2004**, *6*, 1–11.

Wilhelm, S. W.; Suttle, C.A. Viruses and nutrient cycles in the sea viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **1999**, *49(10)*, 781-788.

Williamson, K. E.; Corzo, K. A.; Drissi, C. L.; Buckingham, J. M.; Thompson, C. P.; Helton, R. R. Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods. *Biol. Fertil. Soils* **2013**, *49*, 857-869.

Zaky, A.; Escobar, I.; Motlagh, A. M.; Gruden, C. Determining the influence of active cells and conditioning layer on early stage biofilm formation using cellulose acetate ultrafiltration membranes. *Desalination* **2012**, *286*, 296-303.

CHAPTER 3

MICROBIOLOGICAL STUDY OF BACTERIOPHAGE INDUCTION IN THE PRESENCE OF CHEMICAL STRESS FACTORS IN ENHANCED BIOLOGICAL PHOSPHORUS REMOVAL (EBPR)[1]

3.1. Abstract

Polyphosphate accumulating organisms (PAOs) are responsible for carrying the enhanced biological phosphorus removal (EBPR). Although the EBPR process is well studied, the failure of EBPR performance at both laboratory and full-scale plants has revealed a lack of knowledge about the ecological and microbiological aspects of EBPR processes. Bacteriophages are viruses that infect bacteria as their sole host. Bacteriophage infection of polyphosphate accumulating organisms (PAOs) has not been considered as a main contributor to biological phosphorus removal upsets. This study examined the effects of different stress factors on the dynamics of bacteriophages and the corresponding effects on the phosphorus removal performance in a lab-scale EBPR system. The results showed that copper (heavy metal), cyanide (toxic chemical), and ciprofloxacin (antibiotic), as three different anthropogenic stress factors, can induce phages integrated onto bacterial genomes (i.e. prophages) in an enriched EBPR

[1] This material provided in this chapter is available at Journal of Water Research, 81, 1-14. Reprinted and adapted with permission from Elsevier.

sequencing batch reactor, resulting in a decrease in the polyphosphate kinase gene *ppk1* clades copy number, phosphorus accumulation capacity, and phosphorus removal performance. This study opens opportunities for further research on the effects of bacteriophages in nutrient cycles both in controlled systems such as wastewater treatment plants and natural ecosystems.

3.2. Introduction

Bacteriophages are viruses that are very specific to bacteria as their host (Labrie et al., 2010). Bacteriophages are the most numerous and diverse biological entities in aquatic environments and on the earth. In fact, the bacteriophage population in many cases outnumbers the bacterial population (Suttle, 2007). Although bacteriophages do not directly participate in any metabolic activity in engineered bioreactors such as in activated sludge systems, they can affect the performance of any bioreactor by directly affecting the prokaryotic community through their well-known infection cycles. The bacteriophages need a host to develop/multiply further. There are two alternative developmental pathways of phages; *lytic* which causes the production of progeny virions, and *lysogenic* which results in the integration of the phage genome into the host chromosome, thus forming lysogens, i.e. host cells bearing integrated phage genomes, also known as prophages. The importance of bacteriophages in determining the diversity and structure of a bacterial community has been a subject of keen interest in freshwater (Middelboe et al., 2008; Wilhelm and Matteson, 2008; Kenzaka et al., 2010), terrestrial (Swanson et al., 2009; Wang et al., 2011; Fancello et al., 2012), and marine (Suttle, 2005; Breitbart, 2012; Rowe et al., 2012) studies with several alarming findings (Weinbauer and Rassoulzadegan, 2004; Suttle, 2005). However, almost no information is available

about the role of bacteriophages in the activated sludge process, especially those infecting prokaryotes participating in nutrient removal.

Bio-P removal is accomplished through a process known as enhanced biological phosphorus removal abbreviated as EBPR and the responsible organisms are widely referred as polyphosphate accumulating organisms (PAOs). Despite several significant advances related to understanding the microbiology and physiology of PAOs, process upsets are still observed in activated sludge for phosphorus removal (Oehmen et al., 2007) and these upsets are often related to various types of negative effects experienced by key bacteria (Goel and Motlagh, 2014). The role of bacteriophages in such upsets has not been considered so far. It is generally accepted that bacteriophages exist wherever bacteria are living. For example, activated sludge systems have been shown to contain 10^8~10^9 phages per ml (Otawa et al., 2007; Brown et al., 2015), a number comparable to or greater than the number of phages found in most of the aquatic systems. This number suggests that there must be significant phage-bacteria interactions in activated sludge bioreactors.

In the lysogenic infection cycle, the integrated prophage has two implications for their bacterial host: *(1) the prophage can stay dormant but can come out of the bacterial DNA (a process known as induction) and enter into the lytic cycle and, (2) the prophage can carry certain special genes and if expressed, these gene/s can provide special genetic "identity" to its host bacterium.* The first one threatens the survival of the bacterial community and the second one provides more diversity (through phage mediated horizontal gene transfer) to the prokaryotes in terms of their capabilities to affect the surrounding environment.

Under certain conditions, a developmental switch consisting of prophage induction can occur under which damage to the host DNA takes place resulting in the induction of prophage and the induced prophage enters into lytic cycle (Wegrzyn and Wegrzyn, 2007). Although the reasons for this developmental stage are not clearly understood, some environmental factors, especially in studied aquatic systems, are believed to cause these switches resulting in significant bacterial lysis (Paul, 2008). Activated sludge processes often face stress factors in the forms of unwanted trace organics (i.e. estrogens and antibiotics), metals, toxins and other chemical changes. The big question is whether these factors are causing prophage induction in activated sludge bacteria leading to cell death. A simple search on several activated sludge bacterial genomes available through public databases using prophage finder tools revealed that many of these bacteria have prophage elements on their genome. This implies that these bacteria are susceptible to prophage induction (i.e. possibility of the development switch).

Recently, we investigated the effect of metal, cyanide, pH, and organic loading on prophage induction in *Nitrosospira multiformis*, an ammonia oxidizing bacterium, and demonstrated that indeed these chemicals can cause prophage induction within certain concentration ranges (Choi et al., 2010). Previously conducted research showed that certain stress factors cause nitrification activity to slow down and reasons were attributed to adverse effects of these stress factors on the nitrifying population (Love and Bott, 2002). Our findings not only complimented the reports by these researchers but also showed the impact of prophage induction on nitrifying bacteria leading to cell lysis and process upset. This was also the first finding correlating nitrification upsets to phage dynamics, albeit using the cultured ammonia oxidizing bacterium.

In this chapter, we present our recent findings related to the effect of selected stress factors on biological phosphorus removal efficiency and phage induction. This research was prompted by the reason that the metagenome of "*Candidatus* Accumulibacter phosphatis", the primary agent for enhanced biological phosphorus removal identified so far (Crocetti et al., 2002), suggests that this uncultured organism was also previously infected by bacteriophages causing community and performance changes in phosphorus removal activated sludge (Barr et al., 2010). Although recent research also shows diversity in this uncultured PAO, *Candidatus* Accumulibacter phosphatis remains undisputed as the main PAO participating in EBPR.

The primary objectives of the current study were to examine the effect of heavy metal, toxic chemical, and an antibiotic on EBPR performance. The research examined whether these stress factors can cause prophage induction in an EBPR reactor community. Unfortunately, the PAO community still remains uncultured although they can be enriched in the laboratory-scale reactors. Hence, our approach enriched PAOs in laboratory-scale reactors and then studied the effect of stress factors in phage induction. To best of our knowledge, no study has ever been carried out to research bacteriophage induction factors and their influences in the EBPR system. The research has broader implications for two reasons; this research provides information on the inhibitory effects of different chemical stress factors on EBPR and, it will illustrate a novel pathway, i.e. phage induction, through which chemical agents can negatively affect PAOs.

3.3. Materials and methods

3.3.1. Reactor operation

A 2-liter laboratory-scale control EBPR reactor was initially inoculated with a mixed liquor sample of biomass from a local treatment facility. In order to minimize synergy and interference between different stress factors with varying concentrations, an identical test reactor was also initiated, enriched, and used to subject to different stress factors. After the application of each stress factor, the test reactor was restarted with seed biomass from the main control reactor and allowed to reach a steady state in terms of phosphorus removal and/or TSS/VSS ratio. The test reactor was subjected to the stress factors when the EBPR performance was identical to the control reactor. Both reactors were operated on a 6-hr cycle with four cycles per day. Each cycle consisted of a 2 hr anaerobic period, 3 hr aerobic period, and a final 1 hr settling and decanting period. Dissolved oxygen was kept below detection limits during the anaerobic period by purging nitrogen gas and it was maintained at 3-4 mg/l during the aerobic period. In order to mimic a full-scale treatment plant with a lab-scale reactor, the hydraulic retention time (HRT) and solids retention time (SRT) were adjusted to be 18 hr and 10 days, respectively. Readily biodegradable organic carbon in the form of acetate and other micronutrients were pumped into each reactor rapidly at the beginning of each cycle over a period of 5 min. during the anaerobic phase. The feed composition was taken from Goel and Noguera (2006) with the COD/P ratio of 35. An automatic pH controller was used to control the reactor pH to 7.8±0.2. In addition, allylthiourea was added to inhibit ammonia oxidation during the process (Oehmen et al., 2005). The amount of chemical oxygen demand (COD), dissolved phosphorus, total suspended solids (TSS), and volatile suspended solids (VSS) were measured using the Standard Methods (APHA, 1998).

Specific phosphorus uptake and release was measured every 20 min throughout the anaerobic and aerobic cycle and normalized to the amount of volatile suspended solids.

3.3.2. Construction of clone library using ppk1 gene and phylogenetic analysis

A phylogenetic tree was constructed with biomass sample collected from the control reactor at steady state. Because, the experimental test reactor was always seeded with biomass from the control reactor, it is reasonable to assume that the community in the test reactor was similar to that present in the control reactor. In order to construct an overview representation of the phylogenetic tree in the stabilized control EBPR reactor, amplification of *Candidatus* Accumulibacter *ppk1* fragments was carried out on genomic DNA extracted from the EBPR. The 100-µl reaction mixture contained 50 µl of 2× GoTaq PCR buffer (Promega, WI), 2 µl of template DNA with 70 ng/µl concentration, and 400 nM of each forward and reverse primer as detailed in He et al. (2007). The PCR was conducted on a Gradient Mastercycler (Eppendorf, NY) with a program consisting of an initial denaturation step of 10 min at 95°C, followed by 15 cycles of 95°C for 45 sec, 62°C for 1 min, and 72°C for 2 min; and then 20 cycles of 95°C for 45 sec, 58°C for 1 min, and 72°C for 2 min; and a final extension step at 72°C for 5 min. The amplified *ppk1* fragments were purified from 1% agarose gel using a QIAEX II gel extraction kit (Qiagen, CA) according to the manufacturer's instructions. The purified DNA was then ligated and cloned using a TOPO TA cloning kit (Invitrogen, CA). A total of 96 clones were randomly picked followed by plasmid DNA extraction using a Zippy kit (Zymo Research, CA) and subsequently sequenced using ABI 3130 DNA sequencer (Applied Biosystems, CA) at the DNA University of Utah Core Facility. Chimeric sequences were

checked on the sequences prior to the phylogenetic analysis using DECIPHER (Wright et al., 2012). The phylogenetic tree was constructed using a maximum likelihood method with 500 bootstrap replications using MEGA software version 6 (Tamura et al., 2013), and preliminary operational taxonomic units (OTUs) were defined based on a 99% sequence homology between retrieved sequences.

In addition, the abundance of PAOs in biomass samples was determined by 4-6-Diamidino-2-henylindole (DAPI). Following mechanically dispersing the biomass flocs by repetitively passing the sample through a 26 gauge needle, samples were fixed at room temperature with 4% paraformaldehyde. For intracellular polyphosphate visualization, fixed cells were collected by filtering through a 0.22-µm pore-size polycarbonate filter (Fisher Scientific, PA) and transferred on to a gelatin-coated slide before staining. Transferred cells were stained with 5 µg/ml DAPI solution for 30 min and analyzed under BX51 epifluorescence microscopy (Olympus, Japan).

3.3.3. Quantification of ppk1 genes

At the end of the aerobic cycle following spiking with each stress factor, genomic DNA was extracted with a PowerSoil DNA isolation kit (MoBio, CA). As polyP kinase (PPK) is the responsible enzyme for catalyzing polyP synthesis, real-time quantitative PCR (qPCR) was used to quantify relative abundance of *ppk1* gene as a genetic marker to target and demonstrate *Candidatus* Accumulibacter lineage diversity in the bacterial community using clade-specific *Candidatus* Accumulibacter *ppk1* primer sets (He et al., 2007). For the primer sets having different annealing temperatures, the amplification and quantification was carried out separately. The qPCR was accomplished in a final volume of 20 µl reaction with 10 µl of 2× SYBR Green Mastermix (Thermo Scientific, MA), 1.0

µl of template DNA equivalent to 5 to 10 ng DNA, 10 µM sense and antisense primers and 2.5 µg BSA to limit the amplification inhibition (Kreader, 1996). The qPCR conditioning consisted of an initialization step of 10 min at 95°C followed by 15 cycles of 45 sec denaturation at 95°C, 1 min annealing at 61°C, and 2 min extension at 72°C; and then 20 cycles of annealing at 58°C and extension at 72°C for 2 min. Triplicate samples along with a negative control were run in each reaction.

3.3.4. Prophage induction using mitomycin C

Mitomycin C is a universal inducer that works by damaging the cellular DNA, activating the SOS repair system (including enzyme RecA), which cleaves a repressor and induces the lytic cycle (Weinbauer and Suttle, 1999). In order to check whether the bacterial community in the enriched EBPR reactors had inducible phages, 15 ml of mitomycin C (1 µg/ml) was added to same volume of mixed liquor sludge. After incubating the mixture overnight at room temperature on a gyratory shaker, phage extraction and enumeration was carried out as mentioned in section 3.3.6. As mentioned, mitomycin C was used as a positive control in a serum bottle only to evaluate the induction potential. The control biomass without adding mitomycin C was also included to compare the VLPs changes with the overnight induction.

3.3.5. Applying different stress factors

Different stress factors were applied to the test reactor for one complete cycle each (2 hr anaerobic, 3 hr aerobic, 1 hr settling and decant) and the effects of the stress load were studied in that cycle. The reason behind applying stress factor over one cycle was to simulate shock load of different chemical stresses such as generally observed in

full scale treatment plant. As wastewater from mines and leachate from municipal landfills have considerable levels of copper pollution, copper was used as a surrogate for heavy metal in the EBPR system. A study on concentrations of copper in the influents of 239 wastewater treatment plants showed a range of 0.0001-36.5 mg/l with the median value of 0.4 mg/l (Minear et al., 1981). Therefore, Cu(II) in different concentrations ranging from 0.05 to 1 mg/l was spiked by dissolving $CuCl_2$ in the EBPR reactor to evaluate the prophage induction potential. Cyanide, as a highly toxic chemical to most microorganisms, is produced by human industrial activities such as organic chemical production and the mining industry. Influent wastewater typically has up to 0.5 mg/l total cyanide, which is usually dissolved and in the form of metal-cyanide complexes (Zheng et al., 2004). Potassium cyanide (KCN) is one of the principal human-made cyanide forms in the environment. Therefore, the effects of a toxic substance were assessed by spiking the EBPR with different concentrations of KCN (50-1000 μg/l). Potassium cyanide ionizes in the wastewater and releases a potassium cation and cyanide anion. Finally, ciprofloxacin, as one of the most widespread antibiotics, has been reported in concentrations between 0.0007 and 0.1245 μg/ml in hospital effluent (Hartmann et al., 1999) and was used in ranges of 0.05 to 0.4 μg/ml as the antibiotic stress factor that is being entered into sewage effluent and treatment processes.

3.3.6. Dynamics of bacteriophage population

Following the application of different stress factors, bacteriophages as virus-like particles (VLPs) were extracted and quantified at the end of each cycle in the test reactor. To enumerate VLPs, 15 ml of mixed-liquor sludge was sampled and incubated with the same volume of 3% (w/v) potassium citrate at 4°C for 45 min on a gyratory shaker to

detach any bacteriophages on the cells (Williamson et al., 2003). Afterward, the sample was centrifuged at 6000×g for 30 min and filtered through a 0.22 μm polyethersulfone filter (Millipore, MA) to remove remaining bacterial debris. DNase treatment was carried out by adding 100 μl of the RQ1 DNase buffer (400 mM Tris-HCl, 100 mM $MgSO_4$, 10 mM $CaCl_2$) and 2 μl of RNase free DNase I reagent (Invitrogen, CA) to 900 μl of the filtered sample and incubating the mixture for 30 min at 37°C. Thereafter, in order to stop the DNase activity, 50 mM EDTA solution was added to the sample and the mixture was incubated for 10 min at 65°C. The DNase treated sample was vacuum filtered through a stack of 13 mm filters consisting of a 0.02 μm Anodisc membrane filter (Whatman, England), a 0.65 μm Durapore hydrophilic membrane filter (Millipore, Ireland), and a glass fiber prefilter (Millipore, Ireland). The Anodisc containing captured virus-like particles (VLPs) was stained with 1× SYBR Gold dye (Invitrogen, CA), incubated for 20 min in the dark, and analyzed under BX51 epifluorescence microscopy (Olympus, Japan) using a FITC filter (Choi et al., 2010). The VLPs were enumerated from 20 random micrographs that were digitally captured at magnification of 1000× with DPI-71 camera.

3.3.7. Infection capability of induced phages on the PAOs

To ensure that the induced phages caused by the stress factors are capable of infecting PAOs, the biomass from the test reactor was subjected to the induced phages resulting from the stress factors. As shown in the following section, various stress factors can induce the PAOs' prophage and it was quantified that different concentrations resulted in having different numbers of induced prophages. The concentration for each stress factor that resulted in the highest number of VLPs was selected to investigate the infection capability of induced phages. The induced prophages were isolated by overnight

centrifugation of the biomass at 6000×g followed by filtration of the supernatant through a 0.22 µm polyethersulfone syringe filter (Millipore, MA) to remove remaining bacterial debris. The isolated phages were used to infect the fresh biomass from the test reactor and the phosphorus uptake and release was monitored in a 6-hr cycle consisting of a 2 hr anaerobic period, a 3 hr aerobic period, and a 1 hr settling period.

3.3.8. Quantification of glycogen accumulating organisms (GAOs)

A set of primers including GAOQ431 (TCCCCGCCTAAAGGGCTT) and GAOQ989 (TTCCCCGGATGTCAAGGC) were used to quantify the glycogen accumulating organisms (Crocetti et al., 2002). The qPCR was carried out in a final volume of 20 µl reaction with 10 µl of 2× SYBR Green Mastermix (Thermo Scientific, MA), 1.0 µl of template DNA, 10 µM forward and reverse primers, and 2.5 µg BSA to limit the amplification inhibition. The qPCR was conducted on a Gradient Mastercycler (Eppendorf, NY) with a program consisting of an initialization step of 3 min at 95°C and 45 cycles of 30 sec at 95°C, 45 sec at 60°C, and 30 sec at 72°C (Winkler et al., 2011).

3.3.9. Transmission electron microscopy (TEM)

The extracted bacteriophages obtained after the application of each stress factor were purified by isopycnic centrifugation on a LB-70M ultracentrifuge (Beckman Coulter, CA) at 115,000×g for 3 hr through a cesium chloride gradient. In order to minimize interference of cesium chloride with electron microscopy, overnight dialysis was carried out through dialysis tubing with a nominal MWCO of 3500 (Fisher Scientific, PA) at 4°C. 10 µl of the purified phage sample was loaded on a 400-mesh formvar carbon-coated copper grid and allowed to settle for 1 min. Subsequently, phages

were negatively stained with 1% (w/v) uranyl acetate for 1 min and excess stain was removed by bibulous paper. To study each sample, three grids were used as triplicate while the grids were allowed to air dry prior to examination with Tecnai T12 transmission electron microscopy (FEI, Japan). The accelerating voltage used for imaging was 120 kV, and all phages' micrographs were recorded at 54000× magnification.

3.3.10. Bacteriophage isolation from full-scale EBPR and PAOs infection

In a full-scale wastewater treatment plant, free floating lytic phages can infect their host bacteria. Hence, in order to evaluate the PAOs infection ability of free floating phages generally present in activated sludge bioreactors, the mixed liquor from a local full-scale EBPR wastewater treatment plant was sampled to isolate the free lytic phages. After incubation of the mixed liquor with the same volume of 3% (w/v) potassium citrate at 4°C for 45 min on a gyratory shaker, the sample was centrifuged at 6000×g and filtered through a 0.22 μm polyethersulfone filter (Millipore, MA) to remove remaining bacterial debris. An Amicon ultrafilter (Millipore, MA) was used to concentrate and purify the phages by centrifugation through a regenerated cellulose membrane of 30,000 MWCO. 20 ml of mixed liquor samples from the lab-scale EBPR reactor was used in serum bottles along with the isolated phages from the full-scale treatment plant. Furthermore, in order to evaluate the phage diffusion inside the cells, PAO flocs were also dispersed using a 26 gauge needle and subjected to phage infection.

3.3.11. Nucleotide sequence accession numbers

The sequences determined in this study using *Candidatus* Accumulibacter phosphatis specific primers have been deposited in the GenBank database under accession numbers KJ769130 to KJ769132 for OTU 1 through 3.

3.4. Results and discussion

3.4.1. Reactor performance

Figure 3.1 shows the performance of the control reactor. As illustrated in this figure, the reactor performance consisted of two phases with a two-fold increase in dissolved phosphorus in the influent after 35 days. The increase in the dissolved phosphorus in the influent was applied to enrich PAOs in the reactor even though the dissolved PO_4-P was higher than the PO_4-P concentrations generally encountered in the influent to municipal wastewater treatment plants.

The average dissolved phosphorus released at the end of the anaerobic cycle was 44.6±2.5 (SD, n=31) mg PO_4-P/l. The dissolved phosphorus in the final effluent at the end of the aerobic cycle was consistently lower than 0.9 mg PO_4-P/l and showed more than 86% PO_4-P removal efficiency.

3.4.2. PAO and GAO ecology in the steady-state EBPR reactor

To understand the ecology of the polyphosphate accumulibacter organisms (PAOs) in the EBPR system, PAO communities were phylogenetically studied using *Candidatus* Accumulibacter phosphatis specific biomarkers (He et al., 2007) to evaluate the phylogenetic distance from their evolutionary origin. In addition, *Candidatus* Accumulibacter phosphatis copy numbers were quantified using qPCR. Based on the

DAPI staining of the biomass taken at the end of cycle (Datta and Goel, 2010), it was concluded that the reactor biomass mostly composed of PAOs in which case, most the DAPI stained biomass showed polyphosphate granules. However, as PAOs are classified as *Candidatus* Accumulibacter phosphatis, they cannot be a single bacterium and in an enrichment state, non-PAOs co-exist with PAOs to contribute of PAOs functioning (Kim et al., 2010; Skennerton et al., 2014).

3.4.3. PAO community using "Candidatus Accumulibacter" specific ppk1 gene biomarkers

After using the *Candidatus* Accumulibacter phosphatis targeted ppk1 specific primers, we were able to retrieve *Candidatus* Accumulibacter *ppk1* genes from the EBPR reactor. Figure 3.2 shows a phylogram of *ppk1* genes from this study illustrating the distribution of *ppk1* clones with respect to other relevant sequences obtained from other studies.

The sequences from other *Candidatus* Accumulibacter enriched lab-scale reactors were also included in the phylogenetic analysis as the reference sequences (He et al., 2007; Kim et al., 2010). The *ppk1* sequences from our EBPR system showed almost 50% of the clones belonged in OTU 2 in which was located in the upper region of the phylogram in clade IIA. In addition, a cluster consisting of OTUs 1 and 3 in clade IA formed a cluster closely related to other uncultured PAOs previously reported in McMahon et al. (2002) and He et al. (2007). Although, the possibility of other ppk1 clades cannot be ruled out because there have been reports (Kim et al., 2010; Skennerton et al., 2014) of a greater diversity of *Candidatus* Accumulibacter phosphatis related PAOs since He et al. (2007) first published their findings, these results demonstrate that

the reactor was enriched in *Candidatus* Accumulibacter phosphatis related PAOs.

3.4.4. Quantification of Candidatus Accumulibacter phosphatis related PAOs using qPCR

As the *ppk1* gene is present in *Candidatus* Accumulibacter in one copy (Martin et al., 2006), its abundance can represent the cell numbers of PAOs, especially those related to *Candidatus* Accumulibacter. In this study, we employed *ppk1* gene specific primers developed by He et al. (2007). Recent reports such as from Skennerton et al., 2014 suggest more diverse nature of *ppk1* gene specific to *Candidatus* Accumulibacter. Hence, the *ppk1* gene primers used in this study may have underestimated the *ppk1* gene copy numbers. Nevertheless, we believe that the *ppk1* gene specific primers we used were able to cover majority of the diversity.

The total *Candidatus* Accumulibacter *ppk1* abundance was calculated as the sum of the *ppk1* abundance detected from five *Candidatus* Accumulibacter clades (IA, IIA, IIB, IIC, and IID). The total *Candidatus* Accumulibacter related PAOs present in the reactor was found to have three dominant clades of IIA (67.3%), IIC (25.6%), and IA (6.5%). Our *ppk1* gene based cloning and sequencing revealed the dominance of PAOs in clade IIA (OTU2) and these results well complimented with our *ppk1* gene based qPCR findings in which case almost 67.3% PAOs belonged to clade IIA. Surprisingly, qPCR results showed PAOs belonging to the IIC clade but the *ppk1* targeted sequencing did not show any OTU in this clade. This finding can be a result of biases involved in the PCR amplification (Suzuki and Giovannoni, 1996) and cloning process and possible preferential ligation (Taylor et al., 2007; Palatinszky et al., 2011) that underrepresent some of the taxonomical groups. Although the phylogenetic tree was constructed from

the control reactor, since the test reactor after each stress factor was seeded from the same source (control reactor), the microbial community was identical to the control reactor.

3.4.5. GAO quantification using qPCR

One possible explanation for the decreased P release and uptake during different chemical stress factors tested is the dominance of glycogen accumulating organisms (GAOs) (Oehmen et al., 2005; Weissbrodt et al., 2013). To examine the presence of GAOs, the DNA extracted from the biomass during different stress factors with various concentrations of each stressor was tested to determine whether GAOs were present as a possible microbial community deteriorating the EBPR performance. The qPCR for all of the samples showed significantly low copy numbers of GAOs in the samples ensuring that GAOs were not present (results not reported). These results demonstrate that the presence of GAOs was not the cause of deteriorated EBPR performance during each stress factor tested.

3.4.6. Prophage induction of PAOs using mitomycin C and prophage characterization

Before any stress factor could be applied to evaluate the possibility of phage induction, prophages were induced and released from the chromosome in the presence of mitomycin C as a universal prophage inducing agent, which is a suitable surrogate for environmental stress factors. It was used in order to confirm the induction process and ensure that the PAOs prophages are capable of being induced and are not in a state of mutational decay (Ra'l and H'bert, 2009). After the overnight incubation of the mixed liquor of the EBPR biomass with mitomycin C, a 10-fold increase in the viral population

was confirmed with SYBR Gold staining. These results suggest that indeed the bacteria in the reactor, which are mostly PAOs, have inducible phages.

To characterize the induced prophages following different stress factors, transmission electron microscopy (TEM) was conducted on the induced phage mixture and TEM micrographs are shown in Figure 3.3. As shown in TEM analysis, induced bacteriophages in the EBPR belong to the family *Podoviridae* with a non-enveloped hexagonal head, icosahedral symmetry and a very short non-contractile tail (Ackermann, 2011). As phages in this family tend to be lytic rather than lysogenic, the decrease in the *ppk1* clades and subsequent decline in the reactor performance following different stress factors confirmed the viruses' significant role and their infection effects in the process as discussed later in this chapter. The presented TEM micrographs were the most redundant morphology found under the electron microscopy compared to the control sample.

3.4.7. Effects of copper as a heavy metal

The effects of copper on the biological phosphorus removal were evaluated for the range of 0.05-1 mg/l of Cu(II). Figure 3.4 shows the results on the effect of added concentrations of Cu(II) on the EBPR potential of the biomass and the number of VLPs. As shown in Figure 3.4(a), lower concentrations of Cu(II) up to 0.1 mg/l did not inhibit the specific phosphorus release and uptake rates. In fact, limited amounts of copper were favorable to phosphorus release and increased the specific phosphorus release rate (SPRR) in the anaerobic phase. As copper is a co-factor for various enzymes and required in trace amounts for the growth and functioning of microorganisms, this might imply that copper ions in low concentrations surrounding PAOs can change the charges of the cell membrane and indirectly affect uptake rate as discussed in a study with Tsai et al (2013).

In fact, Wu and Rodgers (2010) showed that inhibitory effects of copper on EBPR mainly occurred during the aerobic phase. However, at higher concentrations, a decreasing trend in phosphorus release and uptake was recorded. The solid line on the top of bar plots in Figure 3.4(a) shows the percentage of phosphorus removal in the test reactor. As stated before, the percentage of phosphorus removal was not much affected at lower Cu(II) concentrations but a significant decrease in phosphorus removal efficiency was recorded at 0.5 and 1.0 mg/l Cu(II) concentrations.

Figure 3.4(c) shows plots of VLPs counts performed at the beginning of each phase of the cycle during the spiking experiment. The VLPs number increased by the end of the anaerobic phase (i.e. beginning of the aerobic) and the number of VLPs thereafter maintained during the aerobic phase especially at higher copper concentrations. The VLPs results are consistent with the specific phosphorus release and uptake rates presented in Figure 3.4(a). The other notable observation is about the VLPs number. This number which almost tripled as compared to the baseline number, at the 0.5 mg/l Cu(II) spike, whereas the number of VLPs increased only slightly at the 1.0 mg/l Cu(II) spike compared to the baseline. Ideally, the highest increase in VLPs number was expected at the 1.0 mg/l Cu(II) spike, but the VLPs number increase was not as much as it was for the 0.5 mg/l Cu(II) spike. This discrepancy can be explained if we consider the phage inducing capacity of copper versus its direct toxicity to the bacterial cell through an oxidative DNA mechanism. In this case, we hypothesize that the lower concentrations of copper acted as stress agents to cause the prophage induction. However, as shown in previous studies (Menkissoglu and Lindow, 1991; Cabrera et al., 2006) copper at a concentration above 0.9 mg/l exerts a significant inhibitory effect on the cultures and

affected the bacterial population through direct killing of the bacteria by oxidizing the cell wall and the DNA (Dupont et al., 2011) without any phage induction. In the higher concentration of copper, the inhibition mechanism primarily occurs through interaction with the intracellular protein structure and functional groups (Nies, 1999) and a rapid disruption of the cytoplasmic membrane (Sani et al., 2001). In the toxic environment, PAOs consume less energy for PHAs synthesis from VFAs and hence, less poly-P is degraded and less phosphate is released from the cells. This cellular toxicity eventually results in an upset of the entire EBPR. In summary, it can be concluded that Cu(II) acted as phage inducing agent to kill PAOs at lower concentrations and it acted as a biocide through its direct interactions with cell wall, protein, and DNA at higher concentrations. These results are consistent with our earlier findings in which we demonstrated the effect of similar stress factors on phage induction in a nitrifying bacterium (Choi et al., 2010).

Figure 3.4(b) shows different *ppk1* gene clades quantified using quantitative PCR. As presented earlier in this chapter, clades IIA (67.3%), IIC (25.6%), and IA (6.5%) were most abundant in the reactor. It can be seen from Figure 3.4(b) that the relative number of all *ppk1* gene clades showed a decreasing trend with higher copper concentrations. Although the possibility of direct killing of PAOs by copper toxicity especially at higher concentrations cannot be ruled out, the decrease in the *ppk1* copy numbers, especially at lower copper concentrations, suggests that the VLPs counted in the collected sample also came out of the PAOs chromosome, demonstrating that stress factors such as the presence of heavy metals can cause PAOs infection and death through intra-cellular phage induction. In summary, the data related to the specific P release and uptake rates, *ppk1* gene copy numbers and VLPs counts suggest the effect of prophage induction on

PAO community in the presence of lower concentrations of copper.

3.4.8. Effects of cyanide as toxic chemical

The effects of cyanide as a toxic chemical on the EBPR efficiency and phage induction were similar to those obtained by copper spiking. As shown in Figure 3.5(a), low concentrations of cyanide up to 100 μg/l did not show severe adverse effects on specific phosphorus release and uptake rates as compared to the baseline rates, although the overall phosphorus removal efficiency seems to show a decreasing trend (as shown by solid line in Figure 3.5(a)). However, a significant decrease in the specific phosphorus release and uptake rate was observed following the spiking of the EBPR reactor with 500 μg/l and higher of cyanide ion. Figure 3.5(c) shows the VLPs count at the beginning of each phase. Subsequently a nearly 4-fold increase of induced virus particles at 500 μg/l cyanide was recorded at the end of the anaerobic phase (i.e. beginning of the aerobic). The increase in the VLPs count was associated with a remarkable decline in the dominant *ppk1* clades of IIA and IIC shown in Figure 3.5(b).

Although the toxicity of cyanide on PAOs can also result in decreasing efficiency of phosphorus removal at 500 μg/l, it is believed that the significant increase in phage population (VLP numbers) and PAOs infection is the most important reason for the change in the reactor performance. In higher concentration such as at 1000 μg/l, cyanide ion blocks ATP production through the electron transport chain by deactivating certain metabolic enzyme systems particularly cytochrome oxidase, the enzyme involved in the transfer of electrons to oxygen that eventually results in cell death (White et al., 2000).

Possibly, that is the reason why EPBR efficiency and *ppk1* gene copy numbers decreased significantly at 1000 μg/l but the number of VLPs did not increase

correspondingly. For different bacterial cells, EC_{50} (i.e. the concentration of toxic chemical that gives half-maximal response) has been reported in a range of 1-2 mg/l (Marugan et al., 2012). Nevertheless, these results confirm that as in the case of heavy metal, toxic loading can negatively impact PAOs by inducing the prophage at low concentrations and can cause severe toxic effects by interfering with cell components at higher concentrations.

3.4.9. Effect of ciprofloxacin as an antibiotic

The presence of antibiotics in municipal wastewater treatment plants is well known (Karthikeyan and Meyer, 2006). Hence, we examined the effect of ciprofloxacin as a model antibiotic on EBPR performance in order to evaluate whether it also causes EBPR upset through prophage induction. The EBPR performance in terms of specific phosphorus release and uptake against different spiked concentrations of ciprofloxacin is shown in Figure 3.6(a). With increasing concentrations of ciprofloxacin, the specific phosphorus release and uptake decreased. Even the smallest concentration of ciprofloxacin, 0.05 μg/ml, affected EBPR in which the specific phosphorus release decreased from 0.29 to 0.25 PO_4-P/mg VSS day and the specific phosphorus uptake decreased from 0.20 to 0.17 PO_4-P/mg VSS day. The phosphorus removal efficiency, shown by a solid line in Figure 3.6(a) also dropped steadily with increasing concentrations of ciprofloxacin.

As shown in Figure 3.6(c), ciprofloxacin caused an induction in viruses at the lowest concentration and resulted in a decrease of the reactor's overall performance. This induction most affected the removal performance with a dose of 0.1 μg/ml and dramatically decreased the copy number of dominant *ppk1* clades of IA, IIA and IIC

(Figure 3.6(b)). Although lower dosages of ciprofloxacin showed an increase in the bacteriophage population, since there was not a significant decrease in the performance and *ppk1* clades, 0.05 µg/ml of ciprofloxacin might have caused prophage induction in the non-PAO community resulting in the drastic increase in phage counts. In other words, most PAOs were resistant to ciprofloxacin at 0.05 µg/ml and majorly the non-PAOs mixed community was affected and showed phage induction. Also, the prophage induction at 0.1 µg/ml of ciprofloxacin caused the clades diversity to decrease in terms of copy numbers, resulting in an even distribution of clades IA, IIA, IIC, and IID.

Bacteriophage induction due to ciprofloxacin was previously mentioned in studies on *Staphylococcus aureus* (Goerke et al., 2006) and *Escherichia coli* strains (Bielaszewska et al., 2012). However, ciprofloxacin has been shown to have a toxic impact on sludge bacterial growth with a half maximal effective concentration (EC_{50}) between 0.31-1.22 µg/ml (Halling-Sorensen et al., 2000). Hence, in the presence of 0.4 µg/ml of ciprofloxacin, the decrease in bacteriophage population can be attributed to bactericidal effects through the inhibition of DNA gyrase, which is an enzyme that relieves the strain while double-stranded DNA is being unwound (Drlica and Zhao, 1997).

3.4.10. Infection of PAOs with induced phages following various stress factors

As mentioned earlier, 0.5 mg/l of copper, 500 µg/l of cyanide, and 0.1 µg/ml of ciprofloxacin showed the highest number of induced VLPs. Therefore, the PAO biomass was spiked with these concentrations for the induction and subsequent isolation of induced phages. The induced phages were extracted using a 0.22 µm polyethersulfone

filter (Millipore, MA) to remove the bacterial cells and used for PAOs infection. The isolated phages were spiked in the test reactor to evaluate the effects of these phages in the phosphorus uptake and release. As shown in Figure 3.7, the phosphorus uptake and release showed a remarkable decrease when spiked with copper-induced phages. Similarly, the cyanide and ciprofloxacin-induced phages decreased the phosphorus uptake and release of PAOs showing the potential of phages in PAOs infection. Although the induced phages resulting from these stress factors caused the decline in the phosphorus uptake and release, the possible effects of the cyanide and ciprofloxacin on the PAOs themselves by making them more vulnerable to phage induction and the subsequent steps of the viral cycle should be taken into consideration. In other words, as the decrease in the phosphorus uptake, release, and removal rate was lower than the counterpart in the presence of stress factor, it is possible that the stress factors indirectly affect the PAOs by making the host less capable of resistance and more susceptible to the induced phages.

3.4.11. Infection of PAOs with lytic bacteriophages isolated from a full-scale EBPR

As discussed earlier in the introduction section, bacteriophages have two infection cycles: (1) lytic and (2) lysogenic. During the lysogenic cycle, bacteriophages integrate their genetic material into their host bacterial genomic DNA and the resulting element is known as a prophage. The prophage element can come out (i.e. induced) under certain environmental stress factors and kill its host (Choi et al., 2010). By the use of a heavy metal, a toxic chemical, and an antibiotic as stress factors, we proved that indeed these stress factors cause prophage induction from the bacterial community in an enriched

EBPR reactor. In this section, we report the possibility of a direct infection of PAOs by lytic phages.

In general, the isolation of a lytic phage for its host bacteria requires classical agar overlay techniques in which a clear plaque appears on the plate in the case of an active infection. However, this methodology is possible only if the host bacteria exist in cultured/isolated form. We have demonstrated this technique using a pure culture of multidrug resistant bacteria (Bhattacharjee et al., 2015). However, the challenge with PAOs is that while these unique organisms can be enriched in lab-scale reactors, so far the attempts to culture these organisms have failed. To circumvent the problem, we opted to use an indirect route to evaluate the possibility of a lytic infection on the PAOs community. We isolated a mixture of free floating phages from the mixed liquor of a full-scale EBPR plant and used this mixture to spike the ongoing test reactor with an assumption that the phage mixture may contain lytic phages with host PAOs from our ongoing EBPR reactor. These isolated free phages can be already induced due to environmental stress factors in the wastewater and their effects on the reactor performance were evaluated in this section.

Figure 3.8 shows specific phosphorus release and uptake rates for: the un-spiked biomass, the phage spiked undispersed EBPR biomass, and the phage spiked dispersed (i.e. un-flocculated) EBPR biomass. In the last case, biomass granules or flocs were dispersed to minimize any transport barrier to the added phages in reaching its host PAOs in case there were any. Furthermore, phages were added to the biomass to have a phage-host ratio of 1000 for an optimal infection (Kotay et al., 2011). This was calculated based on stained isolated phages and bacterial cells. As it is shown in Figure 3.8, specific

phosphorus release and uptake decreased for the phage spiked dispersed and undispersed biomasses. It is also evident from Figure 3.8 that the effects were more severe in the dispersed biomass batch reactor, demonstrating that some of the added phages could reach out to the PAOs easily and infected them. The corresponding phosphorus removal efficiency also went down to 54% and 33% for the undispersed and dispersed biomasses respectively. In addition, a significant increase in the VLPs quantity (1.2E10 ± 1E9 compared to 2.5E8 ± 1.3E7 in the dispersed control biomass excluding the phages that were initially added) following applying phage to the dispersed biomass confirmed the infection process.

A morphology-based study was chosen in this research to identify the classification of the free phages including lytic and induced phages in the reactor and TEM micrographs showed a morphological diversity of the bacteriophages following spiking the EBPR with isolated phages from full-scale wastewater treatment facility (Figure 3.9). These results are significant and for the first time demonstrate that lytic phages capable of infecting PAOs do exist in full-scale plants and can cause poor EBPR by infecting the PAOs.

3.5. Conclusion

The challenge in studying the effect of chemical stress factors on PAOs is our inability to culture these unique organisms although they can be enriched in the lab up to high concentrations. Under this challenge, apparently, it was difficult to conclude that all VLPs came from PAOs. However, based on qPCR results and DAPI staining, we can comfortably conclude that the reactor was enriched in PAOs and, PAOs contributed to VLPs number during each chemical stress test. Based on the presented results, 0.5 ppm of

copper, 500 µg/l of KCN, and 0.1 µg/ml of ciprofloxacin resulted in the highest prophage induction and coincided with significant decreases in *ppk1* genes and phosphorus removal performances. Prophages harbored in PAOs were induced when exposed to environmental factors including a heavy metal, a toxic chemical, and an antibiotic. Prophage induction in PAOs could be one of the underlying causes for deterioration in phosphorus removal efficiency and EBPR performance under stress. It is well known that wastewater treatment plants are long-term processes, and the limited short-term study of stress factors on PAOs and the subsequent dynamics of bacteriophages cannot reflect all of the influences on biological phosphorus removal. However, it can reveal a greater extent of knowledge about the process of PAOs infection and the deterioration of EBPR following different physical and chemical stress factors. A mixture of free floating phages was isolated from a full scale EBPR plant and the contents of the reactors were infected with this mixture. It was observed that the dispersed biomass was more sensitive to the added phages and demonstrated that, perhaps, some of the added phages were proved to be lytic to PAOs. These results also have tremendous practical applications. For example, these results emphasize the importance of granular activated sludge processes to diminish the effects of external microbial perturbations such as the induction of prophage and lytic phages infection on key microbial community.

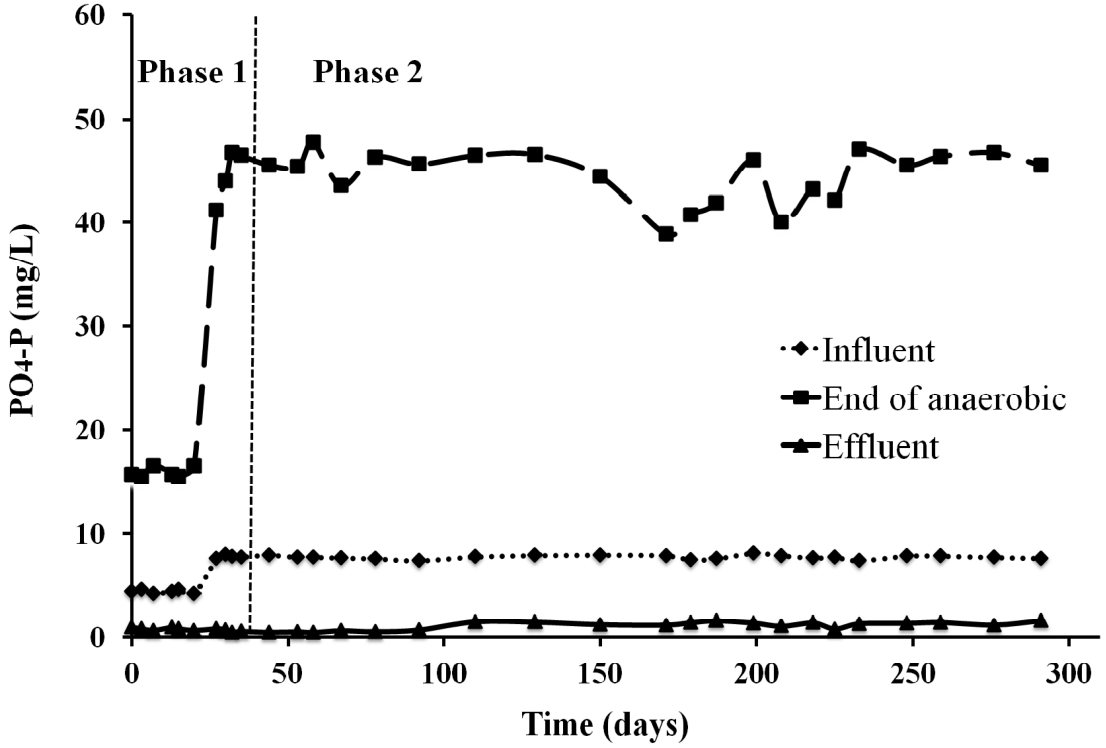

Figure 3.1. Performance of control EBPR. The short dashed line is showing the influent concentration of phosphorus while the solid line is showing the effluent concentration of phosphorus. The long dashed line is showing dissolved phosphorus at the end of the anaerobic cycle

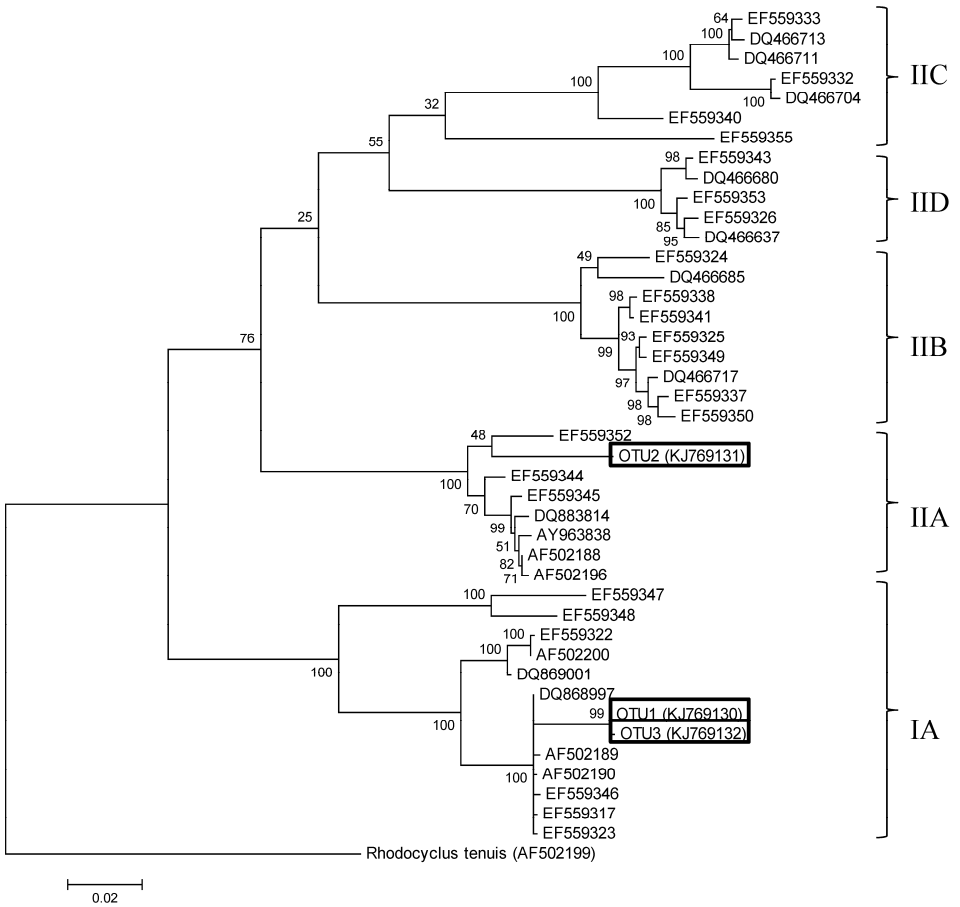

Figure 3.2. Phylogram indicating inferred relatedness of ppk1 genes from the *Candidatus* Accumulibacter lineage. The bordered OTU are the *Candidatus* Accumulibacter sequenced from the control EBPR used in this study. The bar represents 0.02 estimated changes per nucleotide

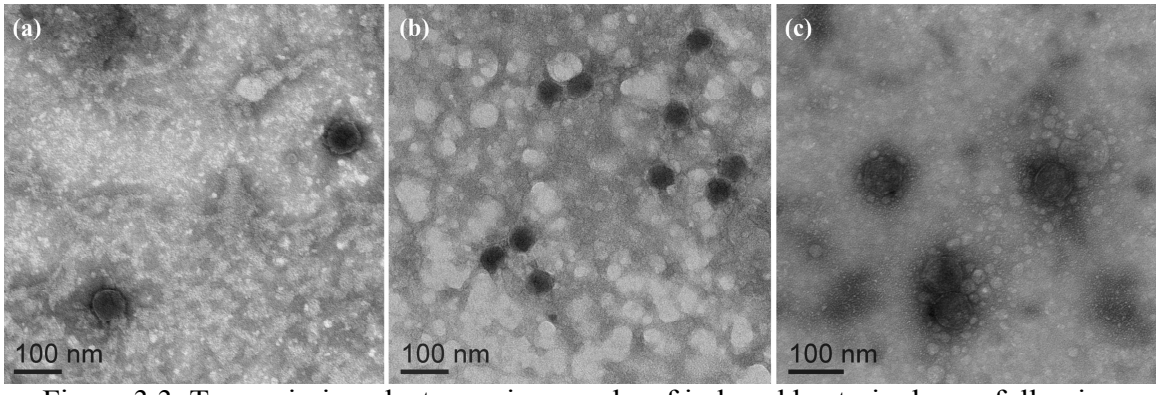

Figure 3.3. Transmission electron micrographs of induced bacteriophages following spiking the EBPR with (a) copper (b) cyanide (c) ciprofloxacin

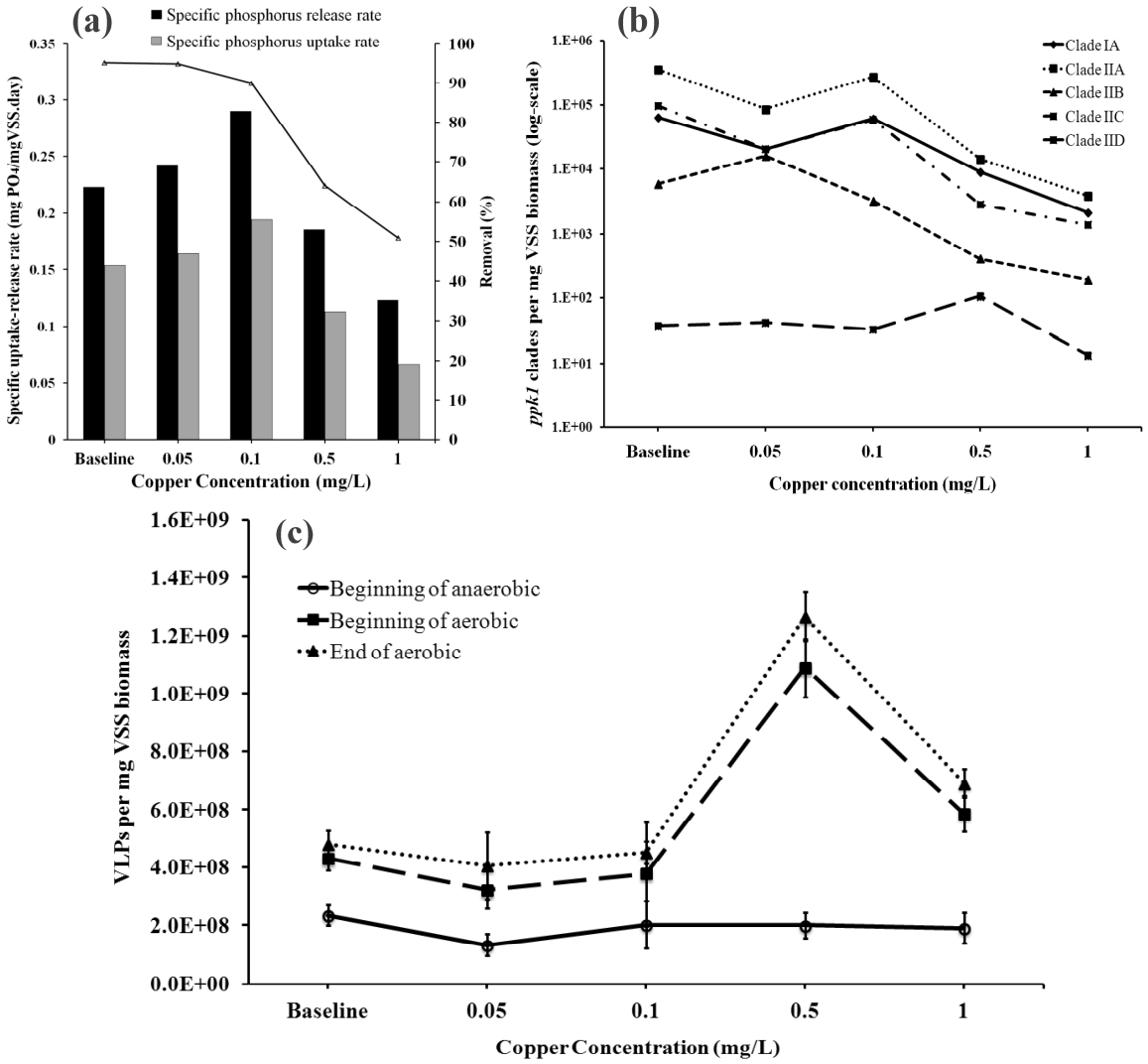

Figure 3.4. Effect of Cu(II) ions in different concentrations on (a) specific phosphorus release and uptake numbers and the corresponding phosphorus removal rate as shown with a solid line on the top of bar plots (b) copy numbers of *ppk1* clades at the end of the aerobic cycle for different concentrations of Cu(II) ions (c) dynamics of induced bacteriophage particles following spiking EBPR with different concentrations of Cu(II) ions at the beginning and end of each cycle

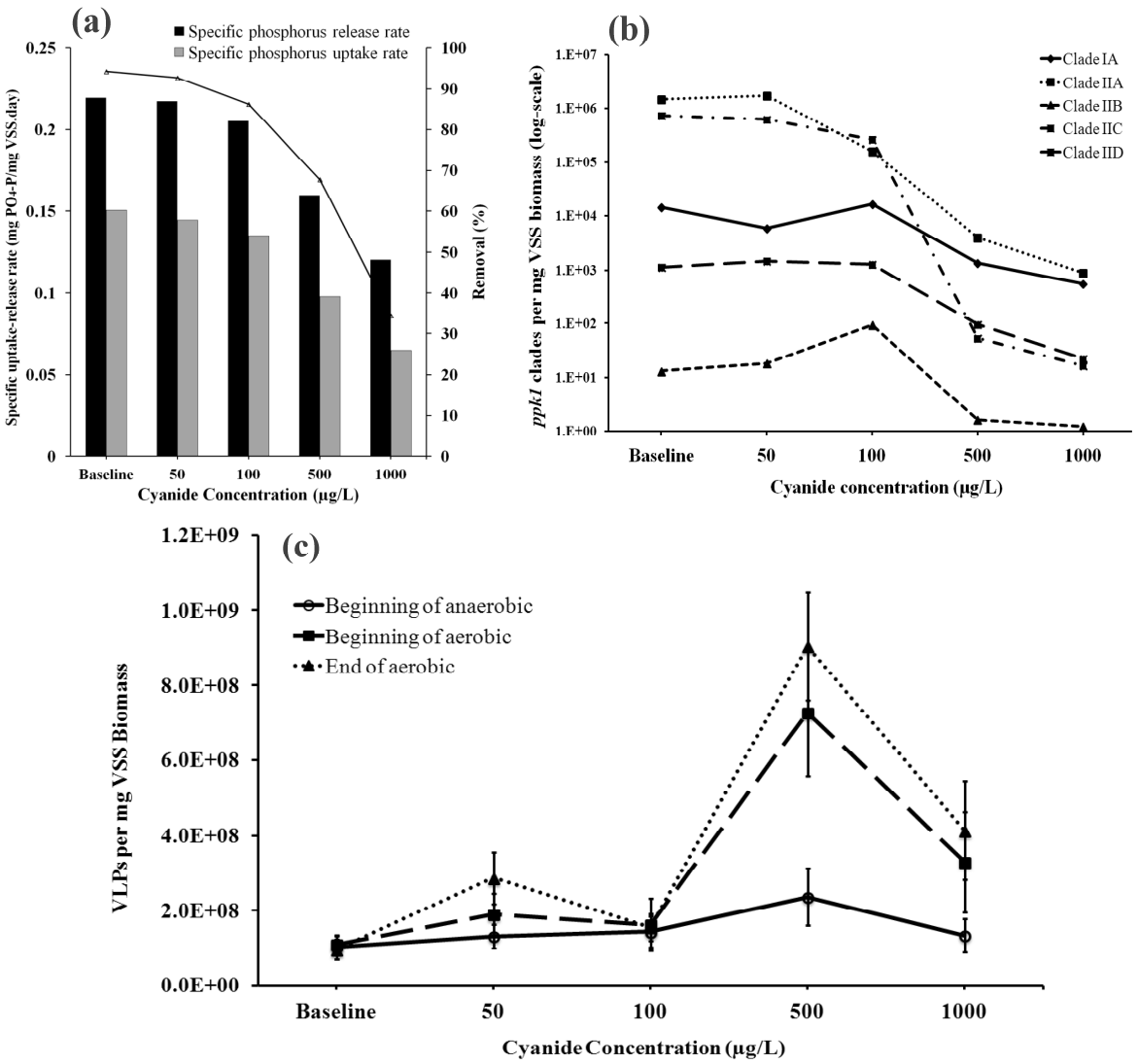

Figure 3.5. Effect of cyanide in different concentrations on (a) specific phosphorus release and uptake numbers and corresponding phosphorus removal rates as shown with a solid line on the top of bar plots (b) copy number of *ppk1* clades at the end of the aerobic cycle for different concentrations of cyanide (c) dynamics of induced bacteriophage particles following spiking the EBPR with different concentrations of cyanide at the beginning and end of each cycle

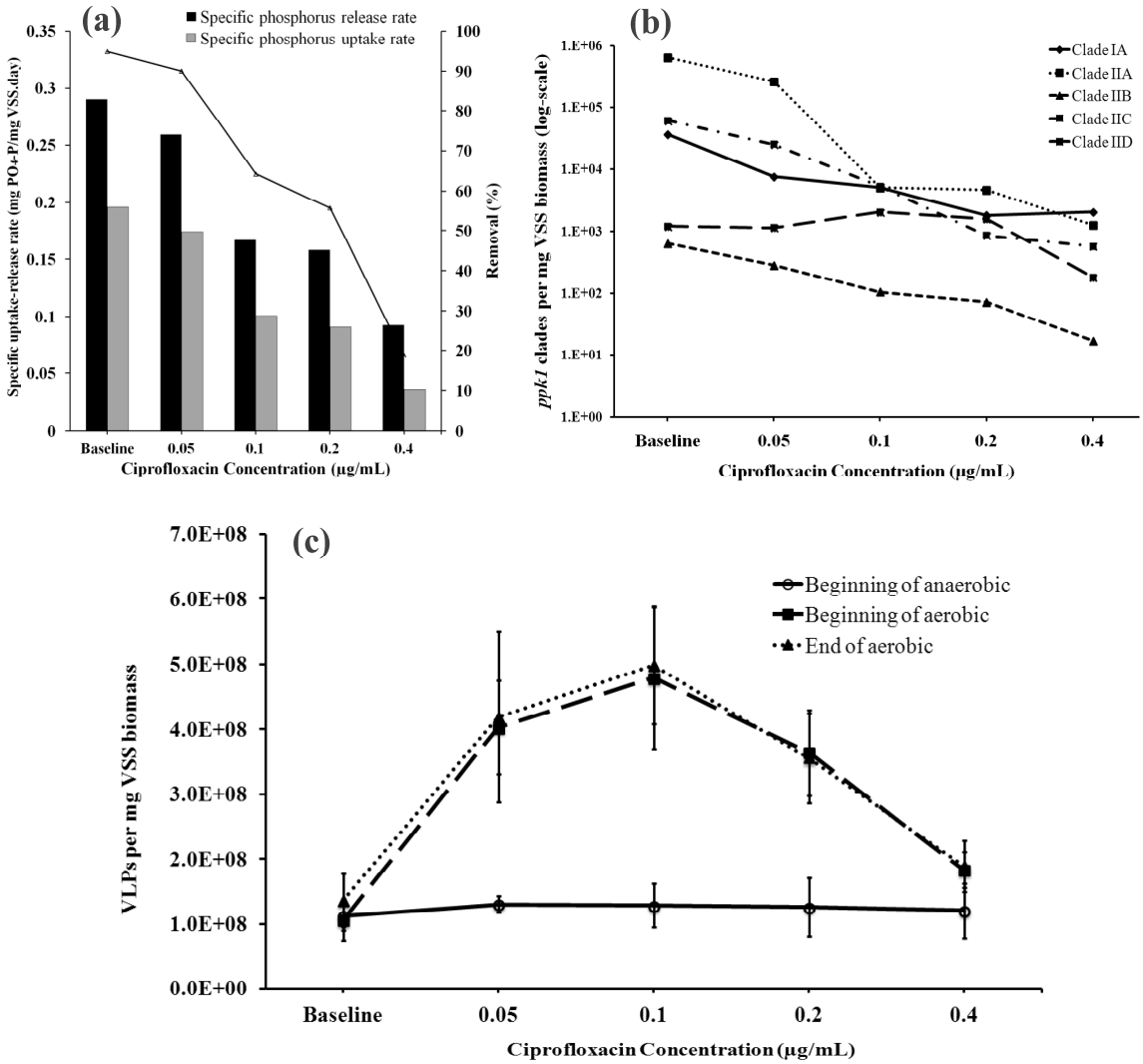

Figure 3.6. Effect of ciprofloxacin in different concentrations on (a) specific phosphorus release and uptake numbers and corresponding phosphorus removal rates as shown with a solid line on the top of bar plots (b) copy number of *ppk1* clades at the end of the aerobic cycle for different concentrations of ciprofloxacin (c) dynamics of induced bacteriophage particles following spiking the EBPR with different concentrations of ciprofloxacin at the beginning and end of each cycle

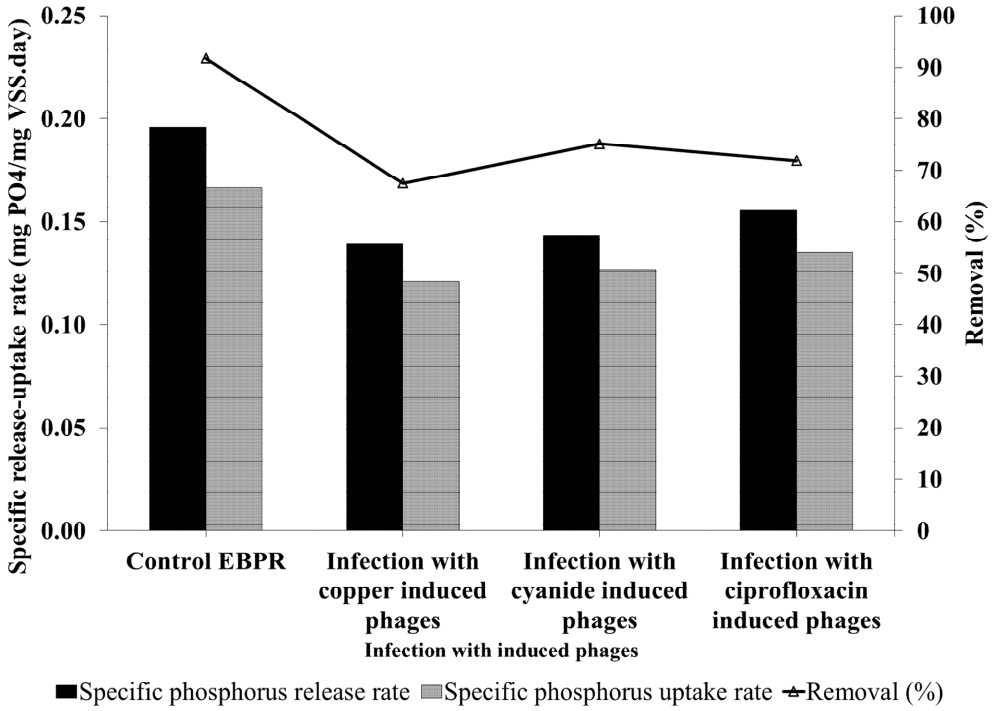

Figure 3.7. The effects of induced phages following different stress factors on phosphorus release and uptake in PAOs

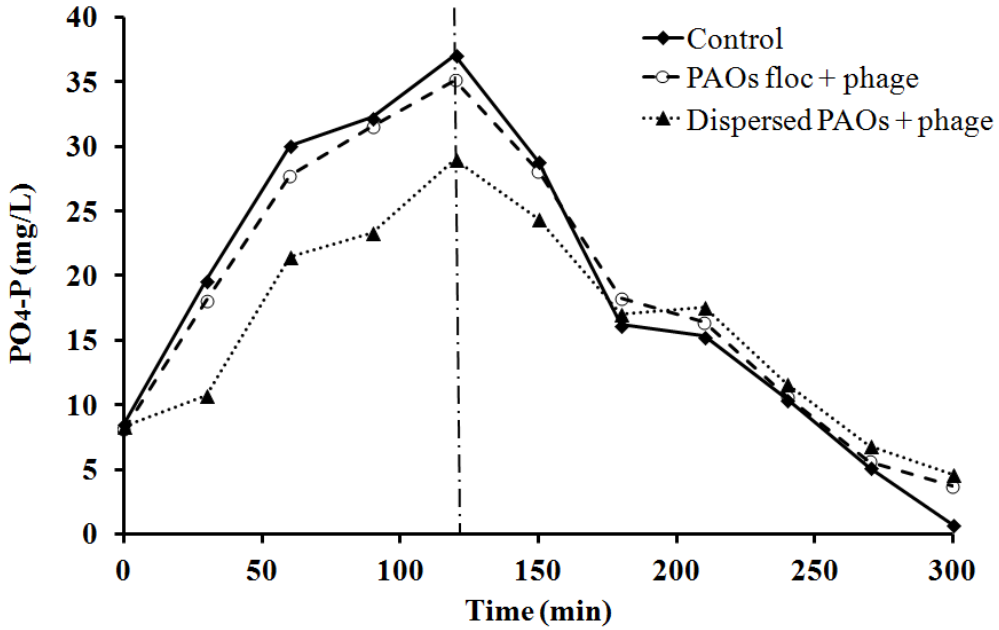

Figure 3.8. Performance of PAOs biomass spiked with isolated phages from full-scale EBPR

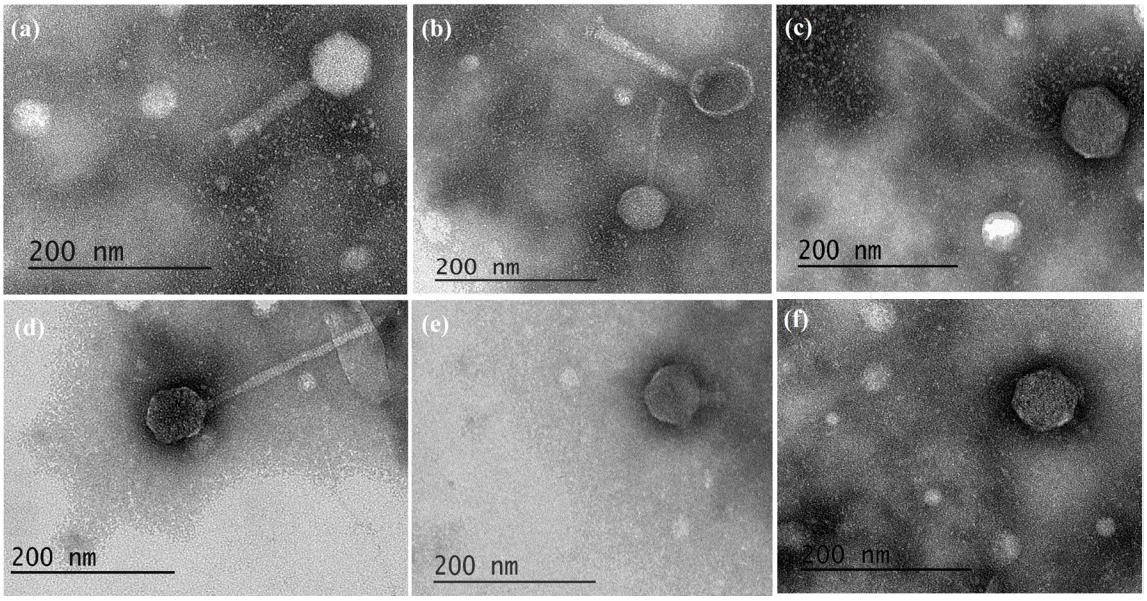

Figure 3.9. Morphological diversity of temperate and induced bacteriophages following spiking the EBPR with isolated phages from a full-scale wastewater treatment facility: (a, b) *Myoviridae* (c, d) *Siphoviridae* (e, f) *Podoviridae*

3.6. References

Ackermann, H. W. Bacteriophage taxonomy. *Microbiol. Aust*. **2011**, *32(2)*, 90-94.

Ahn, K.; Kornberg, A. Polyphosphate kinase from Escherichia coli. Purification and demonstration of a phosphoenzyme intermediate. *J. Biol. Chem*. **1990**, *265(20)*, 11734-11739.

APHA. Standard methods for the examination of water and wastewater; 20th edition, American Public Health Association, Washington, DC. **1998**.

Barr, J. J.; Slater, F. R.; Fukushima, T.; Bond, P. L. Evidence for bacteriophage activity causing community and performance changes in a phosphorus-removal activated sludge. *FEMS Microbiol. Ecol*. **2010**, *74(3)*, 631-642.

Bhattacharjee, A. S.; Choi, J.; Motlagh, A. M.; Mukherji, S. T.; Goel, R. Bacteriophage therapy for membrane biofouling in membrane bioreactors and antibiotic-resistant bacterial biofilms. *Biotechnol. Bioeng*. **2015**, *112(8)*, 1644-1654.

Bielaszewska, M.; Idelevich, E. A.; Zhang, W.; Bauwens, A.; Schaumburg, F.; Mellmann, A.; Peters G.; Karch, H. Effects of antibiotics on Shiga toxin 2 production and bacteriophage induction by epidemic Escherichia coli O104: H4 strain. *J. Antimicrob. Chemo*. **2012**, *56(6)*, 3277-3282.

Breitbart, M. Marine viruses: truth or dare. *Marine Sci*. **2012**, *4*.

Brown, M. R.; Camezuli, S.; Davenport, R. J.; Petelenz-K, E.; Ovreas, L.; Curtis, T. Flow cytometry quantification of viruses in activated sludge process. *Water Res*. **2015**, *68*, 414-422.

Cabrera, G.; Perez, R.; Gomez, J. M.; Abalos, A.; Cantero, D. Toxic effects of dissolved heavy metals on *Desulfovibrio vulgaris* and *Desulfovibrio* sp. Strains. *J. Hazard. Mater*. **2006**, *135(1)*, 40-46.

Choi, J.; Kotay, S. M.; Goel, R. Various physico-chemical stress factors cause prophage induction in *Nitrosospira multiformis* 25196-an ammonia oxidizing bacteria. *Water Res*. **2010**, *44(15)*, 4550-4558.

Crocetti, G. R.; Banfield, J. F.; Keller, J.; Bond, P. L.; Blackall, L. L. Glycogen-accumulating organisms in laboratory-scale and full-scale wastewater treatment processes. *Microbiol*. **2002**, *148 (11)*, 3353-3364.

Drlica, K.; Zhao, X. DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiol. Mol. Biol. Rev*. **1997**, *61(3)*, 377-392.

Dupont, C. L.; Grass, G.; Rensing, C. Copper toxicity and the origin of bacterial

resistance-new insights and applications. *Metallomics* **2011**, *3(11)*, 1109-1118.

Fancello, L.; Trape, S.; Robert, C.; Boyer, M.; Popgeorgiev, N.; Raoult, D.; Desnues, C. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J*. **2012**, *7(2)*, 359-369.

Datta, T.; Goel, R. Evidence and long-term feasibility of enhanced biological phosphorus removal in oxidation-ditch type of aerated-anoxic activated sludge systems. *J. Environ. Eng*. **2010**, *136(11)*, 1237-1247.

Goel, R. K.; Motlagh, A. M. Biological Phosphorus Removal. In: Comprehensive Water Quality and Purification, pp. 150-162, Burlington, MA: Elsevier Science Ltd. **2014**.

Goel R. K.; Noguera D. R. Evaluation of sludge yield and phosphorus removal in a Cannibal solids reduction process, *J. Environ. Eng*. **2006**, *132*, 1331–1337.

Goerke, C.; Koller, J.; Wolz, C. Ciprofloxacin and trimethoprim cause phage induction and virulence modulation in Staphylococcus aureus. *Antimicrob. Agents Chemother*. **2006**, *50(1)*, 171-177.

Halling-Sorensen, B.; Lutzhoft, H. C. H.; Andersen, H. R.; Ingerslev, F. Environmental risk assessment of antibiotics: comparison of mecillinam, trimethoprim and ciprofloxacin. *J. Antimicrob. Chemo*. **2000**, *46*, 53–58.

Hartmann, A.; Golet, E. M.; Gartiser, S.; Alder, A. C.; Koller, T.; Widmer, R. M. Primary DNA damage but not mutagenicity correlates with ciprofloxacin concentrations in German hospital wastewaters. Department of Health and Human Services. **1999**, *36(2)*, 115.

He, S.; Gall, D. L.; McMahon, K. D. "Candidatus Accumulibacter" population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes, *Appl. Environ. Microbiol*. **2007**, *73*, 5865–5874.

Karthikeyan, K. G.; Meyer, M. T. Occurrence of antibiotics in wastewater treatment facilities in Wisconsin, USA. *Sci. Total. Environ*. **2006**, *361(1)*, 196-207.

Kenzaka, T.; Tani, K.; Nasu, M. High-frequency phage-mediated gene transfer in freshwater environments determined at single-cell level. *ISME J*. **2010**, *4(5)*, 648-659.

Kim, J. M.; Lee, H. J.; Kim, S. Y.; Song, J. J.; Park, W.; Jeon, C. O. Analysis of the fine-scale population structure of "Candidatus Accumulibacter phosphatis" in enhanced biological phosphorus removal sludge, using fluorescence in situ hybridization and flow cytometric sorting. *Appl. Environ. Microbiol*. **2010**, *76(12)*, 3825-3835.

Kotay, S. M.; Datta, T.; Choi, J.; Goel, R. Biocontrol of biomass bulking caused by Haliscomenobacter hydrossis using a newly isolated lytic bacteriophage. *Water Res*.

2011, *45(2)*, 694-704.

Kreader, C. A. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Appl. Environ. Microbiol*. **1996**, *62(3)*, 1102-1106.

Labrie, S. J. ; Samson, J. E.; Moineau, S. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol*. **2010**, *8(5)*, 317-327.

Love, N.; Bott, C. Evaluating the role of microbial stress response mechanisms in causing biological treatment system upset. *Water Sci. Technol*. **2002**, *46(1-2)*, 11-18.

Martin, H. G.; Ivanova, N.; Kunin, V.; Warnecke, F.; Barry, K. W.; McHardy, A. C.; Yeates, C.; He, S.; Salamov, A. A.; Szeto, E.; Dalin, E.; Putnam N. H.; Shapiro H. J.; Pangilinan J. L.; Rigoutsos I.; Kyrpides N. C.; Blackall, L. L.; McMahon, K. D.; Hugenholtz, P. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nat. Biotechnol*. **2006**, *24(10)*, 1263-1269.

Marugan, J.; Bru, D.; Pablos, C.; Catala, M. Comparative evaluation of acute toxicity by Vibrio fischeri and fern spore based bioassays in the follow-up of toxic chemicals degradation by photocatalysis. *J. Hazard. Mater*. **2012**, *213*, 117-122.

McMahon, K. D.; Dojka, M. A.; Pace, N. R.; Jenkins, D.; Keasling, J. D. Polyphosphate kinase from activated sludge performing enhanced biological phosphorus removal. *Appl. Environ. Microbiol*. **2002**, *68(10)*, 4971-4978.

Menkissoglu, O.; Lindow, S. E. Relationship of free ionic copper and toxicity of bacteria in solutions of organic compounds. *Phytopathology* **1991**, 81(10), 1258-1263.

Middelboe, M.; Jacquet, S.; Weinbauer, M. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. *Freshwater Biol.* **2008**, *53*, 1069-1075.

Minear, R. A.; Ball, R. O.; Church, R. Data base for influent heavy metals in publicly owned treatment works. US Environmental Protection Agency, Municipal Environmental Research Laboratory. **1981**.

Nies D. H. Microbial heavy metal resistance, *Appl. Microbiol. Biotechnol*. **1999**, *51*, 730–750.

Oehmen, A.; Lemos, P. C.; Carvalho, G.; Yuan, Z.; Keller, J.; Blackall, L. L.; Reis, M. A. Advances in enhanced biological phosphorus removal: from micro to macro scale. *Water Res*. **2007**, *41(11)*, 2271-2300.

Oehmen, A.; Vives, M. T.; Lu, H.; Yuan, Z.; Keller, J. The effect of pH on the competition between polyphosphate accumulating organisms and glycogen-accumulating organisms. *Water Res*. **2005**, *39*, 3727 3737.

Otawa, K.; Lee, S. H.; Yamazoe, A.; Onuki, M.; Satoh, H.; Mino, T. Abundance, diversity, and dynamics of viruses on microorganisms in activated sludge processes. *Microb. Ecol*. **2007**, *53*, 143-152.

Palatinszky, M.; Nikolausz, M.; Svab, D.; Marialigeti, K. Preferential ligation during TA-cloning of multitemplate PCR products—a factor causing bias in microbial community structure analysis. *J. Microbiol. Methods* **2011**, *85(2)*, 131-136.

Paul, J. H. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J*. **2008**, *2(6)*, 579-589.

Ra'l, R. R.; H'bert, E.M. Isolation of phage via induction of lysogens. *In Bacteriophages Humana Press* **2009**, 23-32.

Rowe, J. M.; DeBruyn, J. M.; Poorvin, L.; LeCleir, G. R.; Johnson, Z. I.; Zinser, E. R.; Wilhelm, S. W. Viral and bacterial abundance and production in the Western Pacific Ocean and the relation to other oceanic realms. *FEMS Microbiol. Ecol*. **2012**, *79(2)*, 359-370.

Sani R. K.; Peyton B. M.; Brown, L. T. Copper-induced inhibition of growth of Desulfovibrio desulfuricans G20: assessment of its toxicity and correlation with those of zinc and lead. *Appl. Environ. Microbiol*. **2001**, *67*, 4765–4772.

Skennerton, C. T.; Barr, J. J.; Slater, F. R.; Bond, P. L.; Tyson, G. W. Expanding our view of genomic diversity in Candidatus Accumulibacter clades. *Environ. Microbiol*. **2014**, *17(5)*, 1574-1585.

Suttle C. A. Viruses in the sea. *Nature* **2005**, *437*, 356–361.

Suttle, C. A. Marine viruses-major players in the global ecosystem. *Nature Rev. Microbiol*. **2007**, *5*, 801–812.

Suzuki, M. T.; Giovannoni, S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol*. **1996**, *62(2)*, 625-630.

Swanson, M. M.; Fraser, G.; Daniell, T. J.; Torrance, L.; Gregory, P. J.; Taliansky, M. Viruses in soils: morphological diversity and abundance in the rhizosphere. *Ann. Appl. Biol*. **2009**, *155(1)*, 51-60.

Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*. **2013**, *30(12)*, 2725-2729.

Taylor, D. L.; Herriott, I. C.; Long, J.; O'Neill, K. TOPO TA is A-OK: a test of phylogenetic bias in fungal environmental clone library construction. *Environ. Microbiol*. **2007**, *9(5)*, 1329-1334.

Tsai, Y. P.; Tzeng, H. F.; Lin, J. W.; Lu, M. S.; Lin, J. Y. Verification of enzymes deterioration due to Cu (II) presence in an enhanced biological phosphorus removal system. *Chemosphere* **2013**, *91(5)*, 602-607.

Wang, G.; Asakawa, S.; Kimura, M. Spatial and temporal changes of cyanophage communities in paddy field soils as revealed by the capsid assembly protein gene g20. *FEMS Microbiol. Ecol*. **2011**, *76(2)*, 352-359.

Wegrzyn, G.; Wegrzyn, A. Genetic switches during bacteriophage lambda development. *Prog. Nucleic Acid Res. Mol. Biol*. **2005**, *79*, 1-48.

Weinbauer, M. G.; Rassoulzadegan, F. Are viruses driving microbial diversification and diversity? *Environ. Microbiol*. **2004**, *6(1)*, 1-11.

Weinbauer, M. G.; Suttle, C. A. Lysogeny and prophage induction in coastal and offshore bacterial communities. *Aquat. Microb. Ecol*. **1999**, *18(3)*, 217-225.

Weissbrodt, D. G.; Schneiter, G. S.; Furbringer, J. M.; Holliger, C. Identification of trigger factors selecting for polyphosphate- and glycogen-accumulating organisms in aerobic granular sludge sequencing batch reactors. *Water Res*. **2013**, *47(19)*, 7006-7018.

White, D. M.; Pilon, T. A.; Woolard, C. Biological treatment of cyanide containing wastewater. *Water Res*. **2000**, *34(7)*, 2105-2109.

Wilhelm, S. W.; Matteson, A. R. Freshwater and marine virioplankton: a brief overview of commonalities and differences. *Freshwater Biol*. **2008**, *53*, 1076–1089.

Williamson, K. E.; Wommack, K. E.; Radosevich, M. Sampling natural viral communities from soil for culture-independent analyses. *Appl. Environ. Microbiol*. **2003**, *69(11)*, 6628-6633.

Winkler, M.; Coats, E. R.; Brinkman, C. K. Advancing post-anoxic denitrification for biological nutrient removal. *Water Res*. **2011**, *45*, 6119 -6130.

Wright, E. S.; Yilmaz, L. S.; Noguera, D. R. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol*. **2012**, *78(3)*, 717-725.

Wu, G.; Rodgers, M. Inhibitory effects of copper on enhanced biological phosphorus removal. *Water Sci. Technol*. **2010**, *62*, 1464-1470.

Zheng, A.; Dzombak, D. A.; Luthy, R. G.; Kavanaugh, M. C.; Deeb, R. A. The occurrence of cyanide formation in six full-scale publicly owned treatment works. *Water Environ. Res*. **2004**, *76(2)*, 101-112.

CHAPTER 4

A METAGENOMIC APPROACH TO STUDY PHAGE-BACTERIAL INTERACTION IN SEDIMENTS OF HYPERSALINE ENVIRONMENT OF GREAT SALT LAKE

4.1. Abstract

Bacteriophages as the most abundant biological entities on Earth pose significant predation pressure over their hosts. This pressure plays a critical role in evolution, diversity, and abundance of bacteria. In addition, phages modulate the genetic diversity of prokaryotic communities through the transfer of auxiliary metabolic genes. Various studies have been conducted in various ecosystems to understand phage-host interactions and their effects on prokaryote metabolism and community composition. However, hypersaline environments remain among the least studied ecosystems at the molecular level and the interaction between the phages and prokaryotes at those habitats is poorly understood. Therefore, we fill this gap of knowledge by analyzing bacteriophage-host interaction in the Great Salt Lake, the largest prehistoric hypersaline lake in the Western Hemisphere. Our in-depth metagenomic analysis allowed us to comprehensively characterize interactions between phage and prokaryotic communities of the Great Salt Lake and determine how these interactions affect the diversity and function of these communities and their contributions to biogeochemical cycles.

4.2. Introduction

Microbial communities, particularly prokaryotes, provide diverse ecosystem functions at aquatic habitats such as nutrient cycling (Kouki et al., 2009; Wang et al., 2013; Motlagh and Goel, 2014), carbon cycling (Jasser et al., 2009; Li et al., 2010), mercury transformation (Chavan et al., 2007; Mitchell et al., 2009; Ramasamy et al., 2012), sulfate reduction (Fortin et al., 2000; Bahr et al., 2005; Faulwetter et al., 2013), and many other biological processes. Therefore, studying aquatic microbial communities can provide important information about nutrient removal processes, ecosystem functioning and drivers of biodiversity (Gough and Stahl, 2010; Kim et al., 2011).

Bacteriophages outnumber prokaryotic cells in many ecosystems, posing significant predation pressure on their hosts (Williamson et al., 2013). This pressure plays a critical role in evolution, diversity, and abundance of prokaryotes (Stern and Sorek, 2011). Bacteriophages infect and lyse up to 40% of the prokaryotic population in marine sediments on daily basis (Middelboe, 2008), resulting in decay of cell mass and affecting the carbon pool in sediments, which causes more nutrients to be released in the water column. In addition, bacteriophages can influence the genetic diversity of prokaryotic communities in many different ways. Phages selectively kill their hosts, often in a 'kill the winner' dynamics (Thingstad et al., 2008), in which the most abundant members of a microbial community are the most targeted by phage infection, consequently having their genes removed from the genetic pool of a given habitat.

Phages can also contribute to genetic diversification of prokaryotic communities through transduction, a form of phage mediated horizontal gene transfer (HGT). Through HGT, organisms can acquire exogenous DNA from closely or distantly related lineages and recognition of this process has caused a shift from a diverging 'tree of life' view to a

'web of life' view of evolution (Olendzenski and Gogarten, 2009; Puigbo et al., 2010). Between 1.6 to 32.6% of genes in microbial genomes are estimated to have been acquired by horizontal gene transfer (Koonin et al., 2001), while up to 81% of the genes are estimated to be affected in prokaryotic genomes if the cumulative impact of horizontal gene transfer towards a lineage is considered (Dagan et al., 2008). Transduction has been studied in various natural environments including freshwater (Kenzaka et al., 2010), marine (Jiang and Paul, 1998), plant-associated (Kidambi et al., 1994), and wastewater systems (Del Casale et al., 2011). However, previous focus has been placed on understanding of phage and prokaryote genomes in the transduction process, while the role of transduction for the diversification of prokaryotic genomes and how it affects ecological processes driven by prokaryotes such as nutrient cycles remains poorly understood. This is of special relevance considering previous research that studied the infection susceptibility of marine microbes to phages isolated from soil, marine sediments, and fresh water and showed that phages move and propagate between major biomes, mediating the transfer of DNA between microbes from very different ecosystems (Sano et al., 2004).

An important component in the study of phage-host interaction is the symbiotic state known as lysogeny, i.e. the stable and non-lytic co-existence of the whole viral genome into the prokaryotic host genome. Such quiescent phage genomes are called prophage, and they can be integrated into the chromosome of the host bacterium or exist on plasmids. Prophages often account for most of the difference between strains of the same microbial species (Canchaya et al., 2003; Dutilh et al., 2014). Although the integration of temperate phages into the host genome can be beneficial to host and phage

(Yosef et al., 2015), these prophages can often be induced by a wide range of environmental stress factors such as nutrients (McDaniel and Paul, 2005) or heavy metals (Motlagh et al., 2015) leading to lysis of the bacterial cell and phage virion release. Moreover, a prophage can dramatically change the phenotype of the hosts via lysogenic conversion (Paul, 2008). Therefore, studying the presence and role of prophages in natural ecosystems is a critical step toward our understanding of the phage-host interaction.

Studies over the previous decades greatly increased our knowledge of the genetic richness of viruses in aquatic environments. Culture-independent techniques such as restriction fragment length polymorphism (RFLP) (Chen et al., 1996) and PCR-based methods such as denaturing-gradient-gel electrophoresis (DGGE) (Short and Suttle, 1999), pulse-field-gel electrophoresis (PFGE) (Bhattacharjee et al., 2015), and hybridization analyses (Wommack et al., 1999) have provided snapshots of diversity among prokaryotic communities. However, understanding the taxonomic viral diversity in the environment is challenging since, unlike bacteria, viruses do not carry universally conserved genetic elements such as ribosomal RNA of bacteria that could be used as taxonomic markers to identify all viruses (Rohwer and Edwards, 2002). Metagenomic sequencing approaches have enabled exploration of genetic diversity within environmental samples, thus evading limitations associated with conventional culture dependent microbiology methods. In addition, by targeting total nucleic acids, metagenomics allows functional and taxonomic characterization of samples without the need for a prior knowledge of the microbial types present in the studied environment.

Previous studies have demonstrated the effect of phages and their infection

processes on the sediment microbiome in different environments such as wetland (Jackson and Jackson, 2008), marine (Rohwer and Thurber, 2009), and freshwater environments (Short and Suttle, 2005). However, hypersaline environments are among the least studied ecosystems, and this has resulted in a lack of knowledge at the molecular level about the interaction between their prokaryotic communities and bacteriophages in such habitats. The Great Salt Lake is the largest prehistoric hypersaline lake in the Western Hemisphere, having unique geology with special ecologic (e.g. nutrient cycles, bird habitat) and economic (e.g. source of trace elements, brine shrimp, and mineral industries harvesting) relevance. Therefore, in the present study, we employed metagenomic analysis to analyze interactions between phages and prokaryotes in the Great Salt Lake, to trace the role of bacteriophages in defining the diversity and dynamism among prokaryotic communities and their nutrient cycles.

4.3. Material and methods

4.3.1. Site description and sediment sampling

Great Salt Lake (GSL) is a terminal lake that represents a hypersaline complex ecosystem with salinity gradient (6%-28%) contaminated with mercury, selenium, nutrients and other contaminants (Williams, 2002). The lake is fed by Weber, Bear, and Jordan rivers, while these rivers carry more than 1.1 million tons of salts annually into the lake, which made GSL as one of the most saline inland bodies of water in the world. The total dissolved mineral accumulation in the lake basin is estimated at 5 billion tons, mainly sodium and chloride, though sulfate, magnesium, and potassium are also abundant (Stephens, 1998). As shown in Figure 4.1, the sediment sample was collected 6 miles west of Antelope Island (40°53'51.3" N, 112°21'00" W) in the South Arm of Great Salt

Lake. Physicochemical data in the water column were obtained using a sounder equipped with pH, specific conductivity, water temperature, depth, and dissolve oxygen sensors. The sampling was carried out in June 2014 with collaboration of United States Geological Survey (USGS). The sediment samples were collected from deep brine layer in 27-ft depth at the surface of the sediment with a stainless steel box corer (Wildco, FL). The water in the interface of brine layer and the sediment was also pumped up and stored in a separate polypropylene bottle. The sediment and water samples were immediately shipped to the laboratory on ice for analysis. The water samples were filtered through 0.45 μm cellulose acetate membrane syringe filter (Millipore, MA) and frozen at -20°C until further mixing with the sediment samples for various analyses.

4.3.2. Nutrients and heavy metal analysis

For nutrient determination, sediment samples were centrifuged at 2000×g for 10 min to extract interstitial and pore water. The pore water was filtered with 0.45 μm mixed cellulose hydrophilic filter paper (Millipore, MA) and ammonia (NH_3), nitrite (NO_2), nitrate (NO_3^-), total nitrogen (TN), total phosphorus (TP), and total organic carbon (TOC) were quantified using HACH methods according to the manufacture instruction. In order to ensure that the high salinity of the samples does not interfere with the measurement procedure and reading with the spectrophotometer, control samples with known concentration of ammonia, nitrite, nitrate, and phosphorus were also used along with each sample. Total solids (TS) and volatile solids (VS) were also measured by using 30 g of sediment sample, heating at 103°C for 24 hr followed by igniting at 550°C for 2 hr according to EPA Method 1684. In addition, to study the biogeochemistry of the sediment sample, trace metal concentrations (Mn, Fe, Co, Ni, Cu, Zn, Se, Mo, and Pb)

were measured by using 15 g of sediment, centrifuging at 5000×g for 10 min and filtering the supernatant with 0.22 µm hydrophilic filter, and analyzed using inductively coupled plasma mass spectrometry (ICP-MS, Agilent 7500ce, CA).

4.3.3. Sediment bacterial nucleic acid extraction

Five grams of sediment was resuspended in sterilized MilliQ water, vortexed for 5 min to homogenize the sediment. The homogenized biomass was then centrifuged for 5 min at 2000×g and 250 µl of clear supernatant devoid any cells was discarded. The pellet was resuspended and 500 µl of biomass was used for DNA extraction with different methods including phenol-chloroform DNA extraction (Kochl et al., 2005), centyl trimethyl ammonium bromid (CTAB) extraction (Zhou et al., 1996), PowerSoil® DNA isolation kit (MoBio laboratories, CA), and PowerMax™ Soil DNA isolation kit (MoBio laboratories, CA) following the manufacturer procedures. In order to minimize the extraction biases, the extracted DNA was pooled and the quantity and quality were verified on NanoDrop ND 2000 spectrophotometer (Thermo Scientific, USA) at 260 and 280 nm. The DNA purity and quantity were also verified on 1.2% agarose gel prior to high-throughput sequencing.

4.3.4. Bacteriophage isolation and nucleic acid extraction

Free phages were extracted with resuspending 250 g of sediment in 3× volume of filter sterilized 1% (w/v) potassium citrate buffer (10 g/l potassium citrate, 1.44 g/l of disodium phosphate, 0.24 g/l of monopotassium phosphate, pH 7). The sediment with potassium citrate buffer was placed on a shaker on ice overnight to bring all the free phages in solution. Thereafter, the mixture was centrifuged at 8000 rpm on Avanti J-E

centrifuge (Beckman Coulter, CA) for 30 min to precipitate the bacterial debris followed by centrifugation of the supernatant at 9000 rpm for 12 hr to concentrate the phages. Following overnight centrifugation, the pellet was resuspended in SMG buffer (5.8 g/l sodium chloride, 2 g/l magnesium sulfate, 5 ml/l of 5% (w/v) gelatin, 50 ml/l of 1M Tris–Cl, pH 7.5) and filtered with 0.22 µm pore size filter paper (Millipore Co., MA) to remove any residual bacterial cell and sediment debris.

The phage particles were purified using cesium chloride (CsCl) gradients at 1.35-1.6 g/ml density by isopycnic centrifugation at 35,000 rpm for 3 hr (Angly et al., 2006). The purification process was carried out twice to ensure the removal of any residual bacterial cell debris and guarantee the purity of viral particles. Before the lysis process in the DNA extraction, the sample was subjected to DNase treatment with DNase I, RNase-free (Thermo Scientific, CA) at 37°C for 30 min. This step removed any residual bacterial DNA and digested free DNAs in the sample followed by phage DNA extraction that was based on the spin column purification using a phage DNA isolation kit (Norgen Biotek Corp., Canada). In addition, to confirm the absence of microbial DNA contamination in the phage DNA extracts, PCR amplification of the hypervariable V4-V9 region was tested using 515F/1492R universal 16S rRNA gene primer set (Diemer and Stedman, 2012). Aliquots of the amplification products were electrophoresed in a 1.2% agarose gel stained with ethidium bromide (10 µg/ml), visualized under UV illumination, and was found to be virtually free of microbial DNA.

4.3.5. DNA library preparation for sequencing

Library construction was performed using the Epicentre EpiGnome Methyl-Seq Kit as described below. Briefly, genomic DNA (~50 ng) was heat denatured and

hybridized with oligonucleotides consisting of random hexamers linked to Illumina P5 adapter sequences. Strand replication was accomplished using EpiGnome polymerase. Double-stranded DNA was heat denatured to enable ligation of the EpiGnome Terminal Tagging Oligo which adds Illumina P7 adapter sequence to the 3' end of the replicated strand. Adapter-ligated DNA molecules were enriched by 10 cycles of PCR and the amplified library was subsequently purified using Agencourt AMPure XP beads (Beckman Coulter Genomics, CA). The concentration of the library was measured using the Qubit dsDNA HS Assay (Invitrogen, CA) and an aliquot of the library was resolved on an Agilent 2200 Tape Station using a D1000 assay to define the size distribution of the sequencing library. Libraries were adjusted to a concentration of approximately 10 nM and quantitative PCR was performed using the Kapa Library Quant Kit (Kapa Biosystems, MA) to calculate the molarity of adapter-ligated DNA molecules. The concentration was further adjusted following qPCR to prepare the library for Illumina sequence analysis and the samples were sequenced on an Illumina MiSeq Bench top DNA sequencer (Illumina, CA) with 300-cycles paired-end at Core Facility, HCI, University of Utah.

In many environmental samples, amplification methods such as random amplified shotgun library (RASL), linker-amplified shotgun library (LASL), and multiple displacement amplification (MDA) have been used to amplify single-stranded templates, increase the nucleic acid yield, and obtain a sufficient quantity of DNA for sequencing using a random amplification protocol (Cantalupo et al., 2011; Bibby and Peccia, 2013). However, the random amplification in these protocols leads to quantitative biases (Abulencia et al., 2006; Zhang et al., 2006) and association of phage taxa to bacterial taxa

are difficult to establish. In addition, it has been shown that amplification processes such as MDA amplifies circular DNA more efficiently than linear DNA, which can result in selectively amplification of the genomes of single-stranded DNA (ssDNA) viruses (Kim et al., 2008) and production of artifacts such as chimeras (Lasken and Stockwell, 2007). One of the remarkable aspects in this study was avoiding application of any amplification process to minimize any biases for the phage metagenomics and analyzing of the phageome data.

4.3.6. Metagenomic analysis

Paired-end raw reads were interleaved, quality filtered, and trimmed using CLC Genomics Workbench v.7.0.4 (CLC Bio, Denmark) with threshold of 100 bp as minimum length of read and Phred score of 28. The trimmed reads were *de novo* assembled using CLC Genomics Workbench v.7.0.4 with the following criteria: word size of 20 bp, automatic bubble size of 50 bp, and minimum contig length of 500 bp. Identification of open reading frames (ORFs) and gene prediction were performed using MetaGeneMark v.2.8 (Zhu et al., 2010) followed by gene annotation using RPSBLAST program (Altschul et al., 1990) on clusters of orthologous group (COG) prokaryotic protein database (Tatusov et al., 2000). Statistical over-representation of annotated COG genes and KEGG pathways was determined by pairwise comparisons of each metagenomic sample using Fisher's exact test, with confidence intervals at 99% significance. In addition, the phage and prokaryote predicted genes were compared to SEED subsystems using a maximum e-value of 1E-5, a minimum identity of 60%, and a minimum alignment length of 100 measured in amino acids for protein databases.

GC content of phage and prokaryote contigs were calculated using a perl

command and graphed along with the scaffold length and contig coverage using ggplot package in R v.3.0.1. In addition, the occurrence of each single tetranucleotide in the phage and prokaryote contigs was calculated using JSpecies v.1.2.1 (Richter and Rossello-Mora, 2009).

4.3.7. Analysis of prokaryote metagenome

The taxonomic composition of the prokaryotic metagenome was determined by comparison of the assembled contigs against NCBI non-redundant (nr) nucleotide database using tBLASTx v. 2.2.28 program (Altschul et al., 1990) with 1E-5 e-value cut-off. Analysis of contigs rather than individual raw reads reduced the number of sequences to be analyzed by eliminating redundant sequences (i.e. identical sequences assemble into the same contig). In addition, contigs increases the chances of finding significant matches in the database by increasing sequence length (Wommack et al., 2008). However, as the BLAST analysis was based on the contigs instead of reads, the contig-based annotation was corrected for assembly depth. Therefore, for the reads that were mapped to contigs, every blast hit was multiplied by the average depth of the reads on that contig. Eventually, significant hits to GenBank entries were recorded in a BLAST output file and imported on MEGAN v.5 (Huson et al., 2007) for interpretation by lowest common ancestor (LCA) method. For the MEGAN analysis, default parameters were selected as follows: minimum support of 5, minimum score of 50, top percent of 10, win-score of 0, and minimum complexity of 0.44, were used for each metagenome.

The affinities of the sequences for known metabolic functions were also annotated using BLASTx with cut-off e-value of 1E-5 against SEED subsystems v.2.0 (Overbeck et al., 2005) using MG-RAST server v.3.3 (Meyer et al., 2008) and KEGG metabolic

pathways (Kanehisa et al., 2007). In addition, for understanding the role of prokaryotic communities in nutrient cycles, the identified prokaryotes were compared manually with curated database on functional genes (FunGene) involved in various biogeochemical cycles (Fish et al., 2013).

4.3.8. Analysis of phage metagenome

In contrast to the standard metagenomics approach in which total DNA is sequenced (Cox-Foster et al., 2007; Roossinck et al., 2010), the initial viral particle purification step prior to sequencing ensures that most of the sequencing reads result in the detection of viral nucleic acid. In addition, DNase treatment on the purified extracted phages before phage DNA extraction from virions excludes any bacterial DNA in the sequencing. However, following high-throughput sequencing of phage sample and prior to further metagenomics analysis, the assembled contigs were compared to RefSeq non-redundant bacterial protein using BLASTx with 1E-5 e-value cut-off to exclude any possible bacterial contamination from the phage contigs. Since the analysis of phage metagenome was also performed on contigs instead of reads, contig-based annotation was corrected for assembly depth as mentioned earlier. For phage taxonomic analysis, the filtered phage contigs were compared against NCBI RefSeq viral genome database by using tBLASTx v. 2.2.28 program with e-value of 1E-5 cut-off. The generated BLAST output file was imported on MEGAN v.5 for interpretation by lowest common ancestor (LCA) method using the default parameters as mentioned above.

In addition, the abundance of phages from the Great Salt Lake was explored with respect to metagenomes of marine, freshwater, and lake sediment samples. Raw reads of aquatic viromes available on Metavir were mapped to assembled phage contigs with

length longer than 20 kb using Bowtie v.2 (Langmead and Salzberg, 2012). Contig abundances were obtained by counting the number of reads mapped to each contig and correcting by contig length. The obtained abundance matrix was plotted as a heatmap in which samples and contigs were clustered according to their Euclidean distance using R v.3.0.1.

4.3.9. Prophage identification and analysis

For achieving more in-detail understanding about the phage-host interaction, mobile genetic elements including prophages, plasmids, and transposons were investigated in the prokaryote contigs. Identification of prophage regions within the metagenomic samples was performed to quantitatively predict the prophage abundance and lysis pattern. For finding the possible prophages, following gene prediction in the prokaryote contigs, proteins belong to integrated prophages were detected using prophinder tool (Lima-Mendez et al., 2008) to identify mobile genetic elements including prophage and plasmids. In order to relate the prophages to their host, prokaryote contigs were compared against ACLAME database (Leplae et al., 2006) using BLASTn search with e-value 1E-5 cut-off, and the prophage hit regions were extracted (excluding viruses and plasmids) from the BLAST output results using a python script followed by comparison with prophages found using Prophinder associated with all of the prokaryotic hosts. Finally, the prophages found in the prokaryote contigs with their corresponding hosts were extracted for generating the taxonomic cladogram. In addition, functional annotation for nutrient cycling of the prokaryotic hosts was performed based on KEGG database and used for generating the prophage-host interaction network.

4.3.10. Genetic homology between phage and their host

Phage genome can acquire some prokaryotic genes during an infection event of the host and therefore, homology of phage and prokaryotic genes can indicate phage-host associations. Sequence similarity searches were used to identify genetic homology and predict phage-host relationships. Hence, the assembled bacteria contigs were compared against phage contigs as the database by using tBLASTx v. 2.2.28 program with 1E-5 e-value cut-off.

4.3.11. Clustered regularly interspaced short palindromic repeat (CRISPR)

Clustered regularly interspaced short palindromic repeat (CRISPR) is an antiviral mechanism in archaea and bacteria wherein genomic sequences from the predatory viruses are integrated in the host genome providing immunity to these viruses (Horvath and Barrangou, 2010). Therefore, CRISPR spacer sequences can provide a direct link between viruses and their prokaryotic hosts (Kunin et al., 2008; Heidelberg et al., 2009; Anderson et al., 2011) and CRISPR spacers may be viewed as a database of fragments derived from phage and plasmid genomes. In our study, all CRISPR arrays in the prokaryote contigs with homology to phage contigs were identified using CRISPR Recognition Tool (CRT) v.1.1 (Bland et al., 2007). Putative protospacer targets were also identified using CRISPRTarget (Biswas et al., 2013), following a BLASTn search of the spacer input against ACLAME, GenBank Environmental, GenBank Phage, RefSeq plasmid, and RefSeq viral databases with the default settings.

4.4. Results and discussion

4.4.1. Characteristics of the environmental sample

Following collection of surface sediment sample with a box corer from 27-ft deep brine layer in Great Salt Lake, Utah, the environmental quality parameters, nutrients, and heavy metals of the sediment sample were measured as summarized in Table 4.1. Measurement of nutrients and heavy metals concentration in the sediment can be an indicator of microbial activities involved in different nutrient cycles. Salinity was measured as the total salt concentration, comprised mostly of Na^+ and Cl^- ions. As shown in the table, the measured salinity in Great Salt Lake was 15% (150 g/l), which is approximately 5 times of the average ocean salinity and 300 times of fresh water salinity (Kerr et al., 2003). Significantly low dissolved oxygen in the deep brine layer and high total phosphorus (TP) and total nitrogen (TN) make the water of the Great Salt Lake an extreme environment, which likely selects for a microbiota adapted to high salt concentrations.

Carbon concentration in the Great Salt Lake was measured in terms of total organic carbon (TOC) and it showed considerable available carbon amount of 20.8 $mg.g^{-1}$ of sediment (2%) compared to average TOC content of 0.5% in the deep ocean (Seiter et al., 2004). For the carbon cycle in the sediment, the organic carbon is mainly derived from freshly deposited litter of plants and also decomposed forms such as humus (Kristensen and Holmer, 2001). In addition, bacterial growth and metabolic activities of methanotrophs, methanogens, and photosynthetic phytoplanktons through fixation of CO_2 can contribute to the high concentration of TOC in the sediment sample. As discussed in the following section for prokaryotic community structure, the presence of halophilic and halotolerant methane producing and methanotroph bacteria in Great Salt Lake sediments

studied with our metagenomic analysis confirmed the presence of bacterial carbon source in the sediment. Furthermore, high concentration of nutrients can increase phytoplankton growth and consequently the sedimentation of debris. As a consequence of increased productivity that results in oxygen depletion, the sediment in the anoxic environment of lake will contain a larger amount of organic matter (Lazar et al., 2012).

The considerable amount of ammonia (NH_3) was measured in the sediment sample compared to EPA aquatic life water quality criteria for ammonia (USEPA, 2013), which results in a very low C:N ratio that can be consequences in excretion of excess nitrogen to the sediment. In addition, following the appearance of invertebrates in early Summer in GSL (Wurtsbaugh, 1992), the algae population declines and ammonia notably increases as a result of excretion of uric acid (Meglitsch, 1967) or other nitrogenous wastes produced by brine fly larvae. Another contributing factor to the high concentration of ammonia can be related to appearance of brine shrimp, which increases the ammonia level due to excretion of their waste product. In addition, ammonium (NH_4^+) can be derived from decomposition of organic matter in the water column, or from diffusion of ammonium from the anaerobic sediment layer into the water column. The most likely fate of ammonium in the aerobic water column is nitrification where dominant autotrophic bacteria oxidize ammonium (NH_4^+) to nitrate (NO_3^-).

The levels of trace metals observed at the deep brine layer suggest that this fraction of the sediment may be a hotspot for nutrient cycles. Many enzymatic pathways of the nitrogen and carbon cycle involve metal cofactors which contain iron or molybdenum. In particular, iron is a necessary requirement and an integral component of many enzymes involved in photosynthesis, electron transport, and nutrient acquisition

(Geider and La Roche, 1994). Furthermore, molybdenum is a required component of enzymes involved in nitrogen cycling pathways such as nitrate assimilation and nitrogen fixation (Glass et al., 2015). For instance, nitrogenase, a critical multicomponent enzyme in nitrogen fixation typically consists of an iron and a molybdenum iron containing subunits (Howard and Rees, 1996). As it is discussed in the gene analysis, these critical enzymes involved in various nutrient cycles were found in both prokaryote and phage contigs, suggesting the phage mediated gene transfer process.

4.4.2. Prokaryote and phage contigs assembly

Following bacteria and phage DNA extraction from the sediment sample, DNA library preparation was performed and the DNA samples were sequenced on an Illumina MiSeq Bench top DNA sequencer (Illumina, CA) with 300-cycles paired-end. High-throughput sequencing generated 15.1 million reads for each of the bacteria and phage DNA samples. Following quality and length control, 1.26 million (8.34%) and 1.27 million (8.39%) reads from bacteria and phage samples were removed, respectively. The filtered and interleaved paired-end reads were *de novo* assembled using CLC Genomics Workbench resulting in 40,223 contigs with average length of 1,575 bp, N50 value (i.e. half of the entire assembly was contained in contigs equal to or larger than) 1,573 bp, and average coverage (i.e. number of nucleotides contributing to a portion of the assembly) of 20× for bacteria reads, and 45,689 contigs with average length of 1,595 bp, N50 value of 1,615 bp, and average coverage of 13× for phage reads. Average coverage of the contigs was calculated by multiplying the read quantity by the average read length and dividing by the average contig length.

4.4.3. Prokaryote and phage contigs analysis

Genetic elements such as phages reproduce inside prokaryote cells and use the cell's replication machinery and therefore, most of the phages are expected to have GC content similar to that of their host. However, due to horizontal gene transfer, phages tend to have an average 4% higher AT content compared to their host and therefore, GC content in the phages is generally lower than their hosts (Rocha and Danchin, 2002). Likewise, in our study, GC-content distribution analysis showed prokaryote and phage contigs were in close proximity. As shown in Figure 4.2(a), the bacteria contigs' GC content had a normal distribution of 47±8% with an average of 46.4%, while phage contigs' GC content were normally distributed at 43±10% with average of 42.5%.

Temperate phages and prophages replicate vertically with the prokaryote chromosome and therefore are subject to 'amelioration' towards the oligonucleotide usage profile of the host that they are infecting (Pride et al., 2006). Although virulent phages are not integrated as prophages and continuously are reproduced and transmitted by bacterial replication, they also present a similar trend (Rocha and Danchin, 2002). Therefore, comparison of oligonucleotide usage profiles of bacteria and phage contigs can be used to predict the phage-host associations (Edwards et al., 2016). As shown in Figure 4.2(b), tetranucleotide distribution plotting the bacteria and phage contigs showed a consistent pattern with regression value of 0.98 from plotting both signature occurrences, which can provide a signal for prediction of phage-host interactions.

To understand the phage-host interaction, several criteria including contigs' GC content, length, and their coverage were employed to sort the major assembly pieces. Therefore, these criteria associated with phage contigs were plotted against the bacteria contigs to determine the shared scaffolds with similar GC content and coverage. As

shown in Figure 4.3, several overlapped phage and bacteria contigs with similar GC percentage and coverage were determined suggesting that phages were well-adapted to the codon usage of their prokaryotic hosts.

4.4.4. Prokaryotic community structure and diversity

Microbial communities and specifically prokaryotes in natural ecosystems are involved in nutrient mobilization and regeneration, primary production, energy fluxes, and so investigating its microbial diversity is essential to understand the ecology of Great Salt Lake. Furthermore, the prokaryotic populations in the sediment play key roles in biogeochemical cycles and their infection by phages can affect various nutrient cycles in which these organisms are involved.

For better understanding of the prokaryote diversity, following tBLASTx analysis against NCBI non-redundant (nr) nucleotide database with 1E-5 e-value cut-off, 72.3% of the contigs matched the bacterial domain, while archaea (1%), eukaryotes (6.4%), viruses (13.8%), and unassigned (6.4%) comprised the rest of the contigs. The prokaryotic taxonomic rRNA analysis revealed that *proteobacteria* (46.2%), *firmicutes* (20.3%), *bacteroidetes* (9%), *actinobacteria* (5.3%), *chloroflexi* (5.2%), *cyanobacteria* (2.7%), and *planctomycetes* (1.7%) were the most dominant bacterial phyla present in the sediment sample. As it is illustrated in Figure 4.4, more than 450 different genera were classified in the sediment sample. The taxonomic diversity of prokaryotic communities revealed that most abundant genera of the annotated contigs are belong to taxa involved in various biogeochemical cycles. *Desulfococcus*, the most abundant genus of the GSL microbial metagenome, plays an important role in cycles of sulfur compounds in sea water (Das et al., 2006). *Halanaerobium*, the second most abundant genus in the GSL sediment, is an

obligatory anaerobic halophile while several strains of *Halanaerobium* such as *H. praevalens* were previously isolated from GSL and showed complex organic matter fermentation and production of intermediary metabolites for other trophic groups such as sulfate-reducing and methanogenic bacteria (Ivanova et al., 2011). *Desulfovibrio* and *Desulfatibacillum*, sulfate-reducing anaerobic bacteria genera, were also found to be abundant in the sediment, confirming the anoxic condition of the brine layer sediment. *Streptomyces* was also an abundant genus in the sediment sample, which is an important organism in carbon recycling that has a crucial role in the environment with broad range of metabolic processes such as degradation of the insoluble remains of other organisms including lignocellulose and chitin (Bentley et al., 2002). Proteobacteria of the genus *Rhodobacter* were also found to be abundant in the metagenome, which are extremely metabolically versatile capable of aerobic and anaerobic respiration and anoxygenic photosynthesis when grown anaerobically in the light (Han et al., 2004). The taxonomic structure, population, and diversity of prokaryotes found in the bacterial contigs are also shown in Table 4.2.

4.4.5. Bacteriophage community structure and diversity

As mentioned earlier, the bacteriophages play important roles influencing the prokaryote population and the phenotypic and genotypic diversity of the phage populations are related to the interaction between phages and their host organisms (Cottrell and Suttle, 1995). Understanding of this diversity will provide a tool to characterize the influence of bacteriophages on the population and diversity of prokaryotic communities and contributes to identify phage-host interaction in natural ecosystems. Since unassembled reads in viral metagenome are often characterized by a

majority of unknown sequences that have no homologues in the database (Mokili et al., 2012), assembled reads present in contigs were used for taxonomic characterization to facilitate their annotation. Following BLASTx against RefSeq non-redundant bacterial protein, 3.99% (1,824 contigs) of the phage contigs were excluded from further analysis due to bacterial contamination. Classification of viral metagenome contigs using lowest ancestor analysis (LCA) with the RefSeq viral database, whereas species-specific sequences were assigned to taxa near the leaves of the taxonomic tree and widely conserved sequences were assigned to high-order taxa closer to the root, revealed *Siphoviridae* (32%), *Myoviridae* (24%), and *Podoviridae* (13%) as the most dominant viral families in the sediment sample. In addition to these dominant phage families, other phages such as unclassified dsDNA phages (6%), unclassified archaeal dsDNA virus (6%), and unclassified *Caudovirales* (3%) were among matched hits in the phage contigs.

In addition, regardless of the growing appreciation on the ecological role played by bacteriophages and an increase in phageome information in public databases, only very little is known about their dynamics in natural ecosystems. Hence, it is imperative for the phage researchers to be able to classify these organisms to determine their dynamics. For our study, we developed a high-resolution insight into the viral biogeography by comparing the abundance of the novel unclassified *Caudovirales* isolated from Great Salt Lake sediments with other phage metagenomes obtained from freshwater lakes, lake sediments, and hypersaline marine environments. As shown in Figure 4.5, a heatmap was generated by comparison of putative “complete” phages from Great Salt Lake and the publicly available phage metagenomes of freshwater, lake sediment, and marine environments. In the generated heatmap, cell color represents the

percentage of homology between GSL contigs and other publicly available phage metagenomes. As the Great Salt Lake is a hypersaline lake fed with fresh water, it was interesting to observe from the high-resolution abundance heatmap that these novel unclassified *Caudovirales* Great Salt Lake phages tend to be more abundant in marine environments than in hypersaline lakes, despite GSL being a land-locked hypersaline lake.

4.4.6. Predicted gene abundance and analysis

Gene prediction and ORF identification in bacteria and phage contigs were performed to identify possible events of horizontal gene transfer between the prokaryotic host and their phages. Great Salt Lake metagenome contigs were classified into one of 22 Clusters of Orthologous Groups (COG) functional categories according to best hit classification through RPS-BLAST against COG reference database. Out of the 122,016 bacterial and 102,594 viral predicted genes, 26,319 and 24,495 could be annotated. Figure 4.6 shows the frequency of bacterial and viral genes in various COG classes.

Abbreviations in COG classification are as follows; D, cell cycle control, cell division, chromosome partitioning; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, post-translational modification, protein turnover, and chaperones; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; A, RNA processing and modification; B, chromatin structure and dynamics; J, translation, ribosomal structure and biogenesis; K, transcription; L, replication, recombination and repair; C, energy production and conversion; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and

metabolism; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport, and catabolism; R, general function prediction only; and S, function unknown.

As chromatin forms chromosomes within the nucleus of eukaryotic cells, category B only belongs to eukaryotes and was not detected in our samples. Category A is attributed to RNA processing and it was measured in very negligible match hits (10 and 15 match hits for bacterial and viral genes, respectively). Based on Fisher's exact test, there were significant differences in the functional gene profiles between COG gene categories ($p<0.05$). The over-representation of genes classified into the 'replication, recombination and repair' with 4,341 and 5,022 match hits for bacterial and viral genes, respectively. The most frequent gene among all bacterial predicted genes (occurrence=416), and the second most frequent gene in the viral predicted genes (occurrence=354) was related to replicative DNA helicase (COG0305), an enzyme that participates in initiation and elongation during chromosome replication with unwinding DNA and exhibiting DNA-dependent ATPase activity. In addition, bacterial (314 match hits) and viral (340 match hits) metagenomes showed high representation of DNA modification methylase (COG0863) gene, which is a part of the restriction-modification systems and responsible for producing a species-characteristic methylation pattern that can be used to protect the bacteria from foreign DNA, such as the one borne by bacteriophages.

The COG analysis was a great approach to summarize the functional gene data and show the differences in functional trends of bacteria and phage samples. However, besides COG analysis, the predicted genes were categorized into SEED Subsystems

(Overbeek et al., 2014) through MG-RAST pipeline to identify all of the responsible proteins and their functions in host metabolism. The enzymes were compared against SEED Subsystems using a maximum e-value of 1E-5, a minimum identity of 60%, and a minimum alignment length of 100 measured in amino acids for protein databases. The minimum alignment length of 100 amino acids (equal to 300 bps) was chosen as the criteria to specify the possible functional genes (not only a portion of genes) that are transferred from prokaryotes to the phages. The common enzymes in bacteria contigs which were also found in phage contigs were extracted and shown in the Table 4.3. The presence of these enzymes in the bacteria and phage contigs was suggesting the phage-mediated gene transfer influencing environmental processes.

4.4.7. Prophage identification

Studying and identification of the prophage in the bacteria contigs were one of the important analyses in understanding the phage-host interaction in the natural ecosystem. In the lysogenic infection cycle, the prophage in the prokaryote chromosome can stay dormant but under certain environmental stress conditions, prophages can be induced to enter into lytic cycle and produce progeny virions (Motlagh et al., 2016). Furthermore, prophages integrated into the prokaryote chromosome can carry genes such as photosynthetic genes and phosphorus uptake genes that provide special metabolic features to its host. A total of 69 prophages (33 intact and 36 questionable prophages) were found on the bacteria contigs, while their hosts were assigned based on literatures. From the total prophages that were identified, 37 prophages were associated with their host in the bacteria contigs and were assigned to their prokaryotic host. For the other 32 identified prophages, their prokaryotic host were not found in the bacteria contigs and

therefore excluded from further analysis. The functional annotation for nutrient cycling of the prokaryotic hosts was based on KEGG database. As illustrated in Figure 4.7, the taxonomy of prophage in the bacteria contigs is plotted on the circular cladogram relating the prophages and their prokaryotic hosts involved in different nutrient cycles. The prophage-host interaction network is also tabulated in Table 4.4 showing the prophage, host genera, and the nutrient cycle in which the host is involved.

4.4.8. CRISPRs identification

One of the major limitations of metagenomic analyses is that when bacteria and viruses are sequenced together, it is often difficult to distinguish between the two types of sequences. Even when phages are isolated and sequenced directly, it is nearly impossible to identify the specific hosts from the viral sequence. Without such information, the analysis of the relationship between phage and host is not possible. In order to approach this problem, in this study we analyzed similarities between bacterial CRISPR spacers and the infecting mobile elements to study these interactions in the natural ecosystem.

Initially, in order to associate the CRISPRs in the bacteria contigs with the phages, the bacteria contigs as query were compared with phage contigs as database using BLASTn with E-value 1E-5 cut-off. A total of 6,025 bacteria contigs with more than 70% homology with phage contigs were extracted. Using CRISPR Recognition Tool (CRT), 96 CRISPR arrays were recognized from these extracted bacteria contigs with an average repeat length of 25 bp and spacer length of 40 bp. These identified CRISPR arrays were explored to determine targets of CRISPR RNAs using CRISPRTarget. Interestingly, identified protospacers showed similarity with eight of the marine phage metagenome from TARA database, which were previously shown to have considerable

homology with the phage contigs in the GSL sample. This correlation shows that previous infections of prokaryotic hosts with these phages must have occurred.

4.5. Conclusion

As one of the most unique hypersaline environments on the planet, Great Salt Lake showed a very high diversity in bacteria and bacteriophage species, which was related to various nutrient cycles. Metagenomic analyses of bacteria and phage sequences revealed the relation between the phage and bacteria as their host through their GC contents with respect to the contig coverage, scaffold length, and tetranucleotide frequency. Furthermore, the presence of prophages and their interaction with the prokaryotic host involved in various nutrient cycles showed the susceptibility of Great Salt Lake environment with possible prophage induction due to different environmental stress factors. In addition, gene prediction in the bacteria and phage contigs showed the shared functional genes in the samples, which could have originated from gene transfer process between the bacterial host and phage. Lastly, identification of CRISPRs in the bacterial contigs confirms prevalence of previous infections among the prokaryotic community by phages similar to those in marine environments.

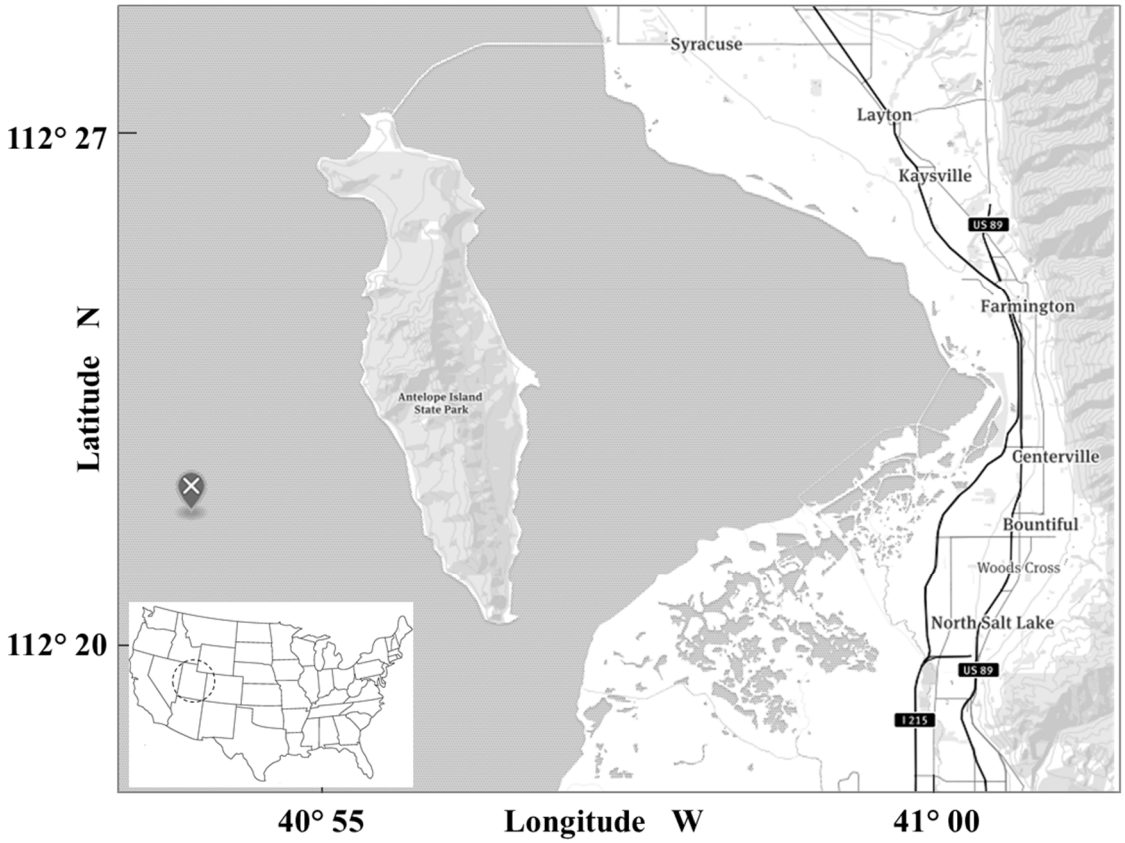

Figure 4.1. The sampling site 6 miles west of Antelope Island in the Great Salt Lake is marked in the satellite map. All of the wetlands as the part of Great Salt Lake are depicted as grey spots in the map

Table 4.1. Environmental quality parameters, nutrients and trace metals present in Great Salt Lake

Specific conductivity (µS/cm)	188,400
Salinity (ppt)	150
pH	7.4
Dissolved oxygen (mg/l)	0.03
Total organic carbon (TOC) (mg/g)	20.8
Total nitrogen (TN) (mg/l)	4.20
NH_3-N (mg/l)	2.29
NO_3^--N (mg/l)	0.04
Total phosphorus (TP) (mg/l)	4.47
Inorganic phosphorus (mg/l)	0.697
Orthophosphate (mg/l)	0.94
Mn (mg/kg)	0.208
Mo (mg/kg)	0.174
Zn (mg/kg)	0.077
Fe (mg/kg)	0.036
Ni (mg/kg)	0.031
Se (mg/kg)	0.010
Co (mg/kg)	0.005
Cu (mg/kg)	0.003
Pb (mg/kg)	0.001

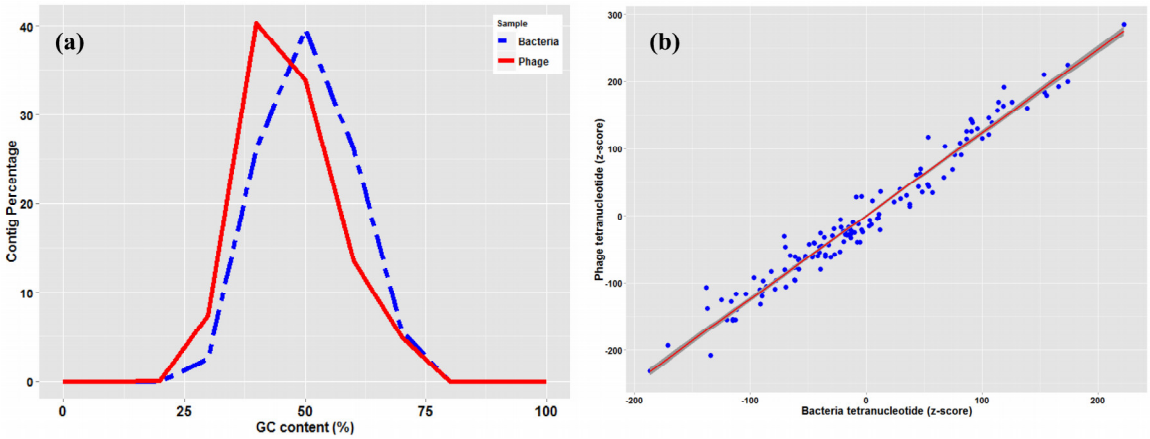

Figure 4.2. Phage-bacterial host interaction (a) GC content distribution of bacteria and phage contigs (b) Standardized (z-score) tetramer word frequencies of phage and bacteria contigs with the regression line with shaded confidence region

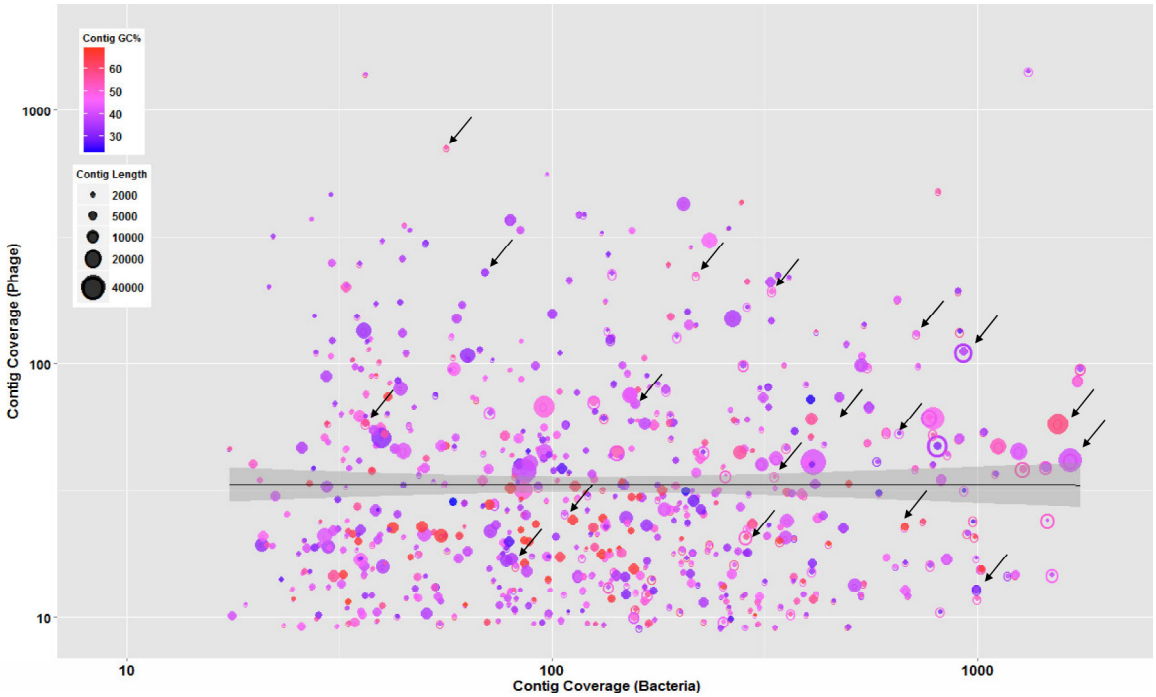

Figure 4.3. Overlapped bacteria (open circles) and phage (solid circle) showing contigs with similar GC content and coverage suggesting phages were well-adapted to the codon usage of their prokaryotic hosts. The regression line with shaded confidence region is shown and arrows are illustrating some of the overlapped contigs

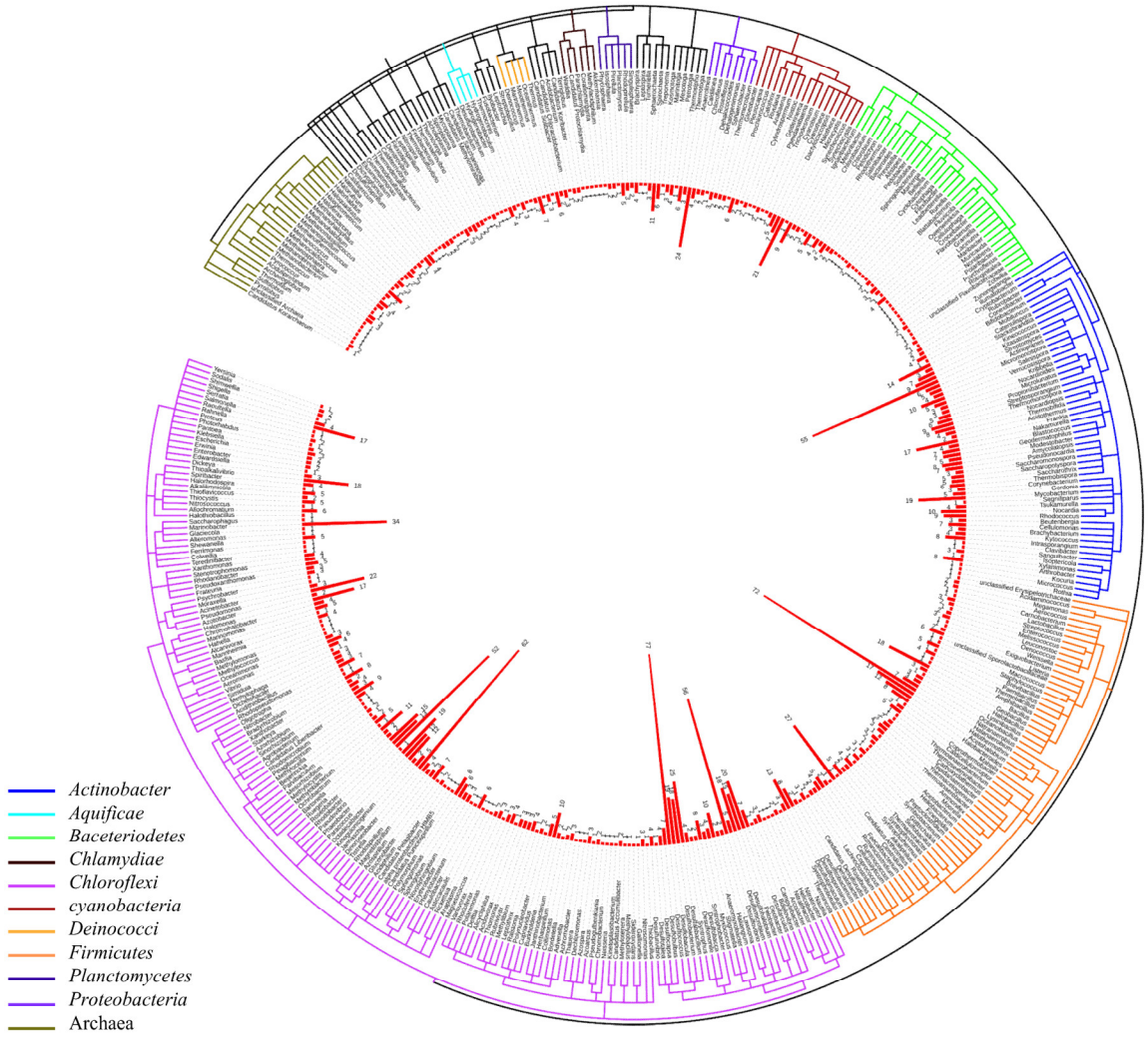

Figure 4.4. Taxonomic cladogram of prokaryotic communities found in metagenomics analysis of Great Salt Lake sediment. The branches are colored according to the phylum to which they belong. Bars display the number of corrected contigs based on the coverage depth assigned to each taxon

Table 4.2. Prokaryote taxonomic population and diversity found in phylum level in the prokaryote contigs

Kingdom	Phylum		Counts
Bacteria	Actinobacteria <phylum>		337
	Aquificae <phylum>		8
	Bacteroidetes/Chlorobi group		61
	Bacteroidetes/Chlorobi group	Bacteroidetes	50
	Bacteroidetes/Chlorobi group	Chlorobi	8
	Bacteroidetes/Chlorobi group	Ignavibacteriae	3
	Caldiserica		1
	Chlamydiae/Verrucomicrobia group		6
	Chlamydiae/Verrucomicrobia group	Chlamydiae	3
	Chlamydiae/Verrucomicrobia group	Verrucomicrobia	3
	Chloroflexi		40
	Chrysiogenetes <phylum>		4
	Cyanobacteria		67
	Deferribacteres <phylum>		5
	Deinococcus-Thermus		18
	Dictyoglomi		4
	Elusimicrobia		1
	Fibrobacteres/Acidobacteria group		10
	Fibrobacteres/Acidobacteria group	Acidobacteria	10
	Firmicutes		314
	Fusobacteria		7
	Gemmatimonadetes		1
	Nitrospirae		5
	Planctomycetes		15
	Proteobacteria		956
	Spirochaetes		26
	Synergistetes		7
	Tenericutes		4
	Thermodesulfobacteria <phylum>		3
	Thermotogae <phylum>		18
	unclassified Bacteria		4
	unclassified Bacteria	candidate division NC10	2
	unclassified Bacteria	*Candidatus* Saccharibacteria	1
	unclassified Bacteria	Thermobaculum	1
Archaea	Crenarchaeota		3
	Euryarchaeota		45
	Korarchaeota		1
	unclassified Archaea		2
	unclassified Archaea	unclassified Archaea (miscellaneous)	2
Not assigned	N/A		119

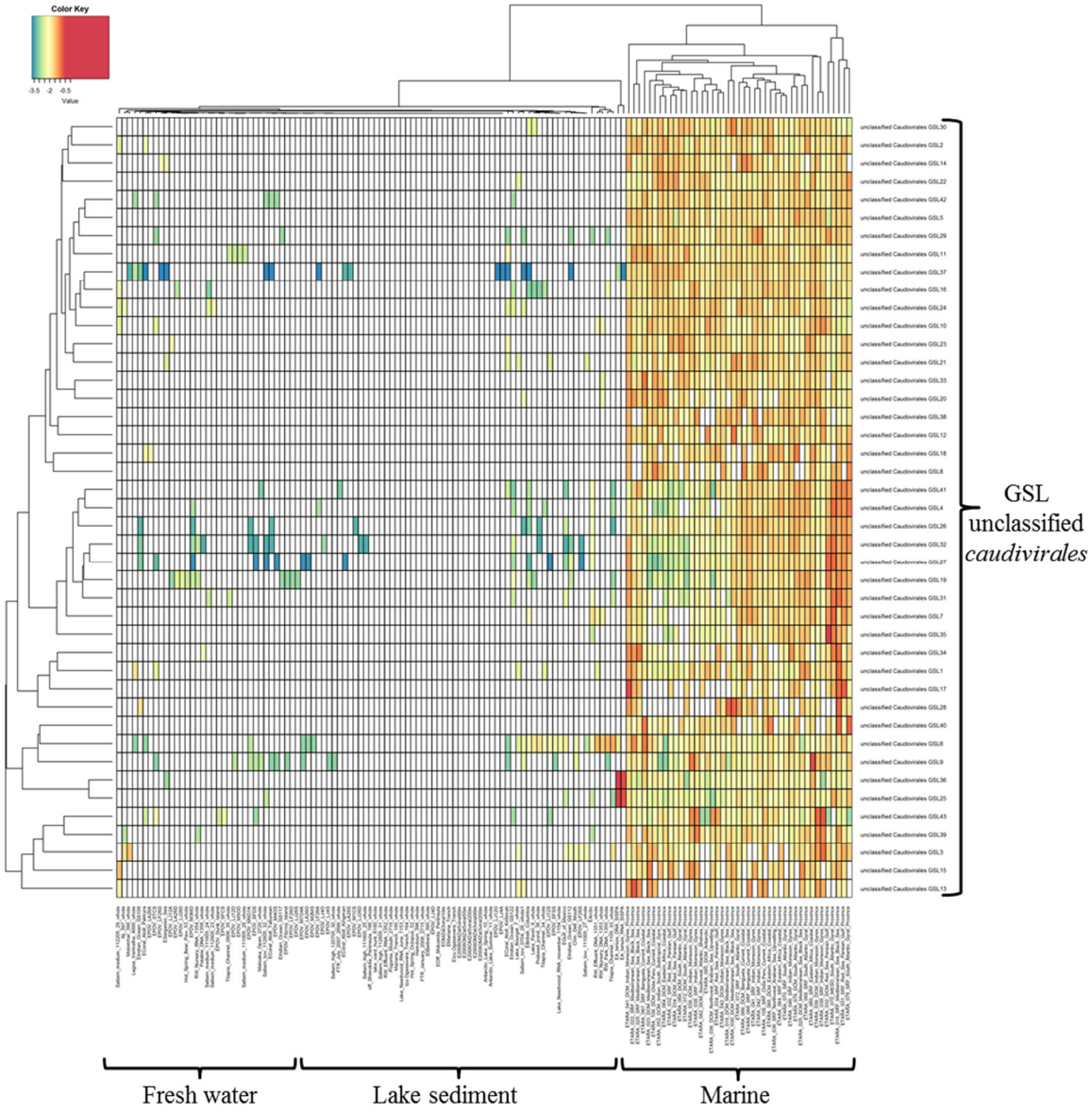

Figure 4.5. Heatmap generated by comparison of putative complete phages from Great Salt Lake and the metagenomes of marine and freshwater lakes publicly available. Cell color represents the relative percentage of identity (log_{10}) of reads in viromes mapped to GSLs phage genome fragments

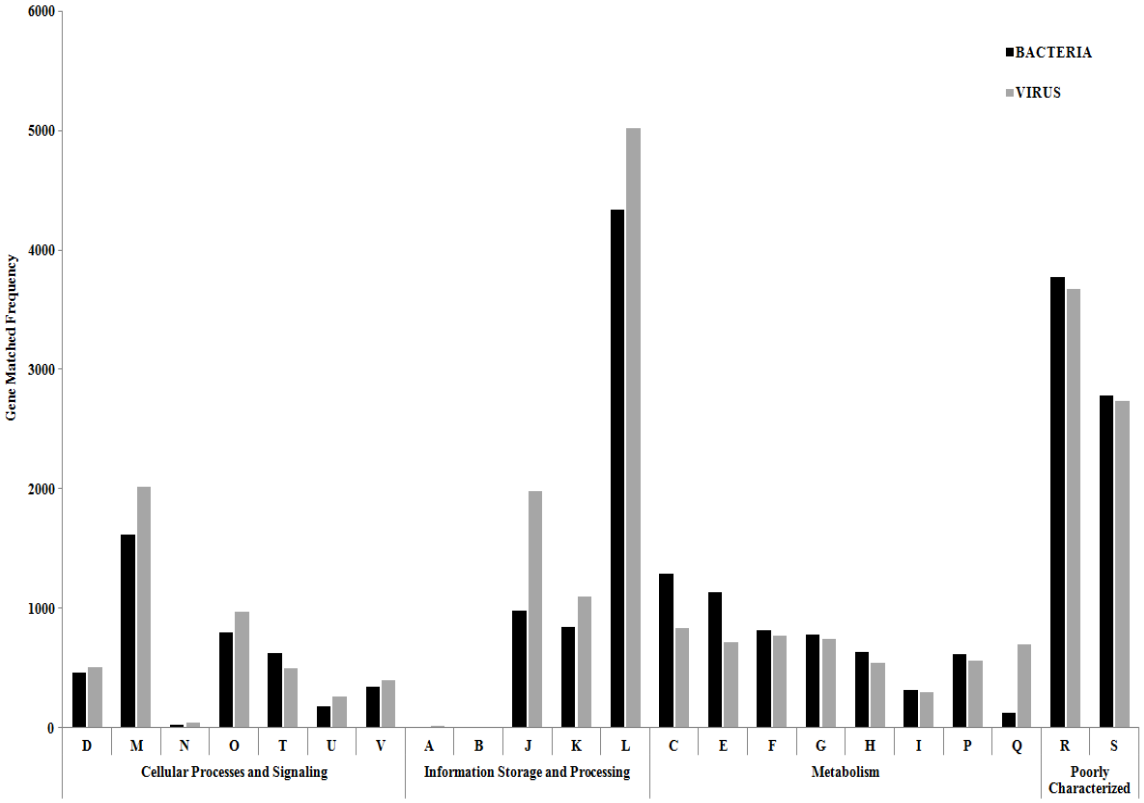

Figure 4.6. Comparison of bacterial and viral functional gene profile classification based on COG classification scheme

Table 4.3. The common functional enzymes in the bacteria and phage contigs

level 1	level 2	level 3	function	Abund	avg % ident	avg align len	# hits
Amino Acids and Derivatives	Aromatic amino acids and derivatives	Aromatic amino acid degradation	Anthranilate dioxygenase reductase	1	92.3	117	1
Carbohydrates	Central carbohydrate metabolism	Dehydrogenase complexes	Enoyl-CoA hydratase [branched-chain amino acid degradation] (EC 4.2.1.17)	32	63.4	139.8	3
Carbohydrates	Organic acids	Methylcitrate cycle	PrpF protein involved in 2-methylcitrate cycle	9	83.9	111.6	8
Carbohydrates	Organic acids	Propionate-CoA to Succinate Module	2-methylcitrate dehydratase FeS dependent (EC 4.2.1.79)	12	83.7	106.2	12
Membrane Transport	-	Tricarboxylate transport system	Ammonia monooxygenase	1	84.6	111	1
Membrane Transport	Protein and nucleoprotein secretion system, Type IV	Vir-like type 4 secretion system	Inner membrane protein forms channel for type IV secretion of T-DNA complex, VirB8	10	94.2	111.4	4
Metabolism of Aromatic Compounds	Peripheral pathways for catabolism of aromatic compounds	Benzoate degradation	Benzoate transport protein	5	68.1	100.2	3
Nitrogen Metabolism	-	Nitrate and nitrite ammonification	Nitrate ABC transporter, permease protein	1	93.3	120	1
Phages, Prophages, Transposable elements, Plasmids	Phages, Prophages	Phage replication	DNA helicase (EC 3.6.1.-), phage-associated	52	69.1	101.2	6

Table 4.3. Continued

level 1	level 2	level 3	function	Abund	avg % ident	avg align len	# hits
Phages, Prophages, Transposable elements, Plasmids	Plasmid related functions	Plasmid-encoded T-DNA transfer	Inner membrane protein forms channel for type IV secretion of T-DNA complex, VirB8	10	94.2	111.4	4
Protein Metabolism	Protein biosynthesis	Translation elongation factors bacterial	Translation elongation factor Tu	1131	82.3	116.7	188
Protein Metabolism	Protein biosynthesis	tRNA aminoacylation, Thr	Threonyl-tRNA synthetase (EC 6.1.1.3), chloroplast	41	62.9	112.1	1
Respiration	Electron accepting reactions	Terminal cytochrome C oxidases	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	221	80.6	101.7	71
Respiration	Electron accepting reactions	Terminal cytochrome d ubiquinol oxidases	putative Cytochrome bd2, subunit I	227	76	104.3	44
Stress Response	Osmotic stress	Choline and Betaine Uptake and Betaine Biosynthesis	Glycine betaine ABC transport system, permease/glycine betaine-binding protein OpuABC	2	70.7	106	1
Stress Response	Oxidative stress	Rubrerythrin	Fe-S oxidoreductase-like protein in Rubrerythrin cluster	2	79.6	100	2

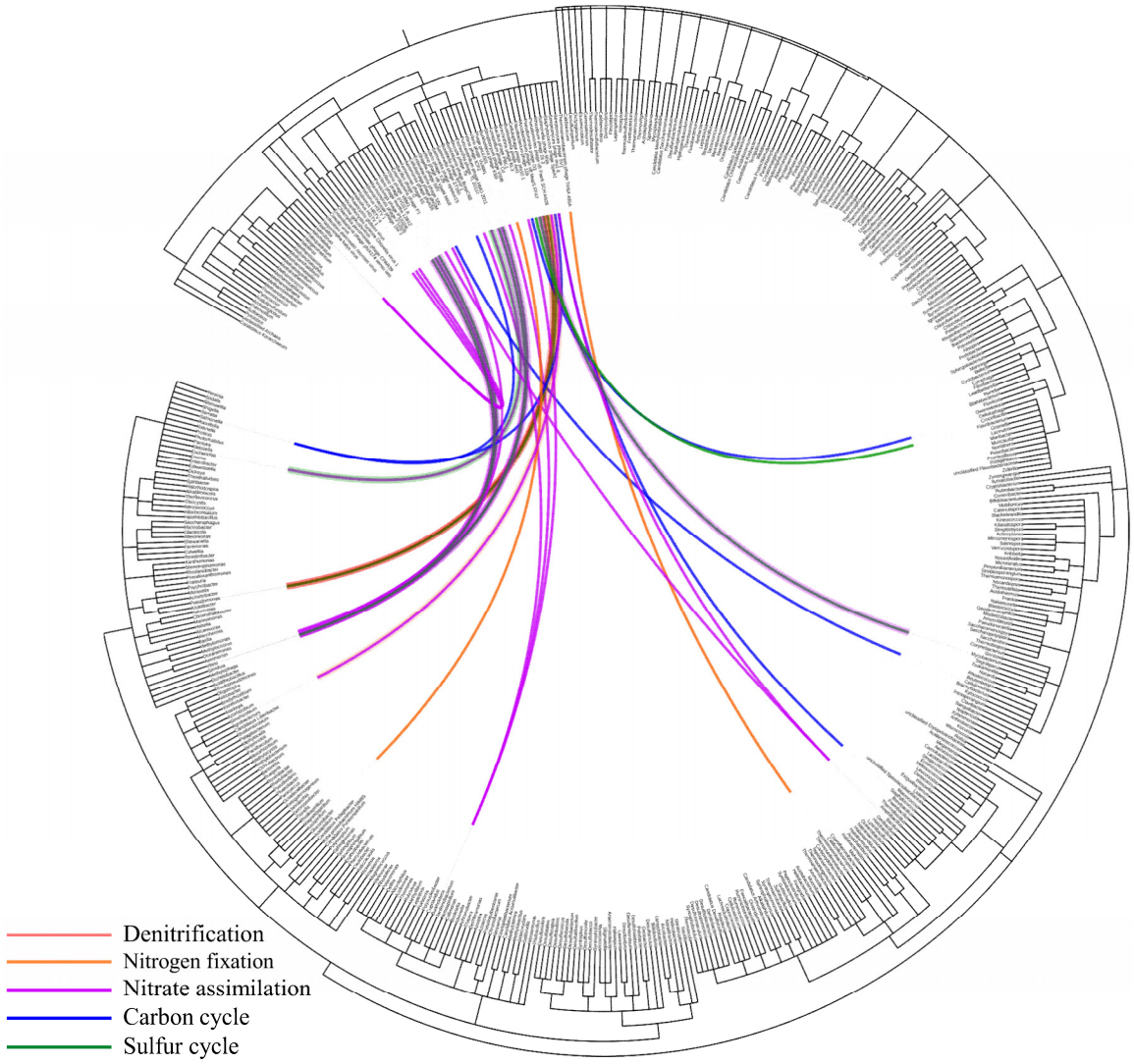

Figure 4.7. Network of prophage and their prokaryotic hosts. The dendrogram is showing the taxonomy of the phage taxa that were found as integrated prophages in the bacteria contigs with respect to their hosts. The network is showing the interaction of the prophages and their prokaryotic hosts and the color of the line indicates the nutrient cycle that these prokaryotic hosts are involved in and susceptibility of these cycles due to presence of infecting temperate phages

Table 4.4. Prophage-host interaction network

Prophage	**Host genera**	**Nutrient cycle involved**
Aeromonas_phage_phiAS5	Aeromonas	Nitrate Assimilation
Bacillus_phage_G	Bacillus	Nitrate Assimilation
Cellulophaga_phage_phiSM	Cellulomonas	Carbon Cycle
Cellulophaga_phage_phi10_1	Cellulophaga	Carbon Cycle
Vibrio_phage_martha_12B12	Vibrio	Sulfur Cycle
Vibrio_phage_martha_12B12	Vibrio	Nitrate Assimilation
Vibrio_phage_pYD21_A	Vibrio	Sulfur Cycle
Vibrio_phage_pYD21_A	Vibrio	Nitrate Assimilation
Vibrio_phage_VBM1	Vibrio	Sulfur Cycle
Vibrio_phage_VBM1	Vibrio	Nitrate Assimilation
Vibrio_phage_ICP2	Vibrio	Sulfur Cycle
Vibrio_phage_ICP2	Vibrio	Nitrate Assimilation
Vibrio_phage_VvAW1	Vibrio	Sulfur Cycle
Vibrio_phage_VvAW1	Vibrio	Nitrate Assimilation
Burkholderia_phage_BcepC6B	Burkholderia	Nitrate Assimilation
Burkholderia_phage_KS9	Burkholderia	Nitrate Assimilation
Burkholderia_phage_AH2	Burkholderia	Nitrate Assimilation
Salmonella_phage_ST160	Salmonella	Carbon Cycle
Salmonella_phage_9NA	Salmonella	Carbon Cycle
Escherichia_phage_TL_2011c	Escherichia	Nitrate Assimilation
Escherichia_phage_TL_2011c	Escherichia	Sulfur Cycle
Azospirillum_phage_Cd	Azospirillum	Nitrogen Fixation
Flavobacterium_phage_11b	Flavobacterium	Sulfur Cycle
Mycobacterium_phage_vB_MapS_FF47	Mycobacterium	Sulfur Cycle
Mycobacterium_phage_vB_MapS_FF47	Mycobacterium	Nitrate Assimilation
Rhizobium_phage_16_3	Sinorhizobium	Nitrogen Fixation
Rhizobium_phage_16_3	Sinorhizobium	Nitrate Assimilation
Staphylococcus_phage_SpaA1	Staphylococcus	Carbon Cycle
Staphylococcus_phage_SpaA1	Bacillus	Nitrate Assimilation
Thermoanaerobacterium_phage_THSA_485 A	Thermoanaerobacterium	Nitrogen Fixation
Pseudomonas_phage_D3	Pseudomonas	Sulfur Cycle
Pseudomonas_phage_D3	Pseudomonas	Denitrification
Pseudomonas_phage_vB_PaeS_SCH_Ab26	Pseudomonas	Sulfur Cycle
Pseudomonas_phage_vB_PaeS_SCH_Ab26	Pseudomonas	Denitrification
Halovirus_HGTV_1	Halorubrum	Nitrate Assimilation
Halovirus_HRTV_4	Halorubrum	Nitrate Assimilation
Halovirus_HSTV_1	Halorubrum	Nitrate Assimilation

4.6. References

Abulencia, C. B.; Wyborski, D. L.; Garcia, J. A.; Podar, M.; Chen, W.; Chang, S. H.; Chang, H. W.; Watson, D.; Brodie, E. L.; Hazen, T. C.; Keller, M. Environmental whole-genome amplification to access microbial populations in contaminated sediments. *Appl. Environ. Microbiol*. **2006**, *72(5)*, 3291-3301.

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol*. **1990**, *215(3)*, 403-410.

Anderson, R. E.; Brazelton, W. J.; Baross, J. A. Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol. Ecol*. **2011**, *77(1)*, 120-133.

Angly, F. E.; Felts, B.; Breitbart, M.; Salamon, P.; Edwards, R. A.; Carlson, C.; Chan, A. M.; Haynes, M.; Kelley, S.; Liu, H.; Mahaffy, J. M. The marine viromes of four oceanic regions. *PLoS Biol*. **2006**, *4(11)*, e368.

Bahr, M.; Crump, B. C.; Klepac-Ceraj, V.; Teske, A.; Sogin, M. L.; Hobbie, J. E. Molecular characterization of sulfate-reducing bacteria in a New England salt marsh. *Environ. Microbiol.* **2005**, *7*, 1175–1185.

Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris. D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3 (2). *Nature* **2002**, *417(6885)*, 141-147.

Bhattacharjee, A. S.; Choi, J.; Motlagh, A. M.; Mukherji, S. T.; Goel, R. Bacteriophage therapy for membrane biofouling in membrane bioreactors and antibiotic-resistant bacterial biofilms. *Biotechnol. Bioeng*. **2015**, *112(8)*, 1644-1654.

Bibby, K.; Peccia, J. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ. Sci. Technol*, **2013**, *47(4)*, 1945-1951.

Biswas, A.; Gagnon, J. N.; Brouns, S. J.; Fineran, P. C.; Brown, C. M. CRISPRTarget: Bioinformatic prediction and analysis of crRNA targets. *RNA Biol.* **2013**, *10 (5)*, 817–827.

Bland, C.; Ramsey, T. L.; Sabree, F.; Lowe, M.; Brown, K.; Kyrpides, N. C.; Hugenholtz, P. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* **2007**, *8(1)*, 209.

Canchaya, C.; Fournous, G.; Chibani-Chennoufi, S.; Dillmann, M. L.; Brussow, H. Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.* **2003**, *6*, 417–424.

Cantalupo, P. G.; Calgua, B.; Zhao, G.; Hundesa, A.; Wier, A. D.; Katz, J. P.; Grabe, M.; Hendrix, R. W.; Girones, R.; Wang, D.; Pipas, J. M. Raw sewage harbors diverse viral

populations. *MBio* **2011**, *2(5)*, e00180-11.

Chavan, P. V.; Dennett, K. E.; Marchand, E. A.; Gustin, M. S. Evaluation of small-scale constructed wetland for water quality and Hg transformation. *J. Hazard. Mater.* **2007**, *149(3)*, 543-547.

Chen, F.; Suttle, C. A.; Short, S. M. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Appl. Environ. Microbiol.* **1996**, *62*, 2869–2874.

Cottrell, M. T.; Suttle, C. A. Genetic Diversity of Algal Viruses Which Lyse the Photosynthetic Picoflagellate Micromonas pusilla (Prasinophyceae). *Appl. Environ. Microbiol.* **1995**, *61(8)*, 3088-3091.

Cox-Foster, D. L.; Conlan, S.; Holmes, E. C.; Palacios, G.; Evans, J. D.; Moran, N.A.; Quan, P. L.; Briese, T.; Hornig, M.; Geiser, D. M.; Martinson, V. A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* **2007**, *318(5848)*, 283-287.

Dagan, T.; Artzy-Randrup, Y.; Martin, W. Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc. Natl. Acad. Sci.* **2008**, *105*, 10039–10044.

Das, S.; Lyla, P. S.; Khan, S. A. Marine microbial diversity and ecology: importance and future perspectives. *Curr. Sci.* **2006**, *90(10)*, 1325-1335.

Del Casale, A.; Flanagan, P. V.; Larkin, M. J.; Allen, C. C. R.; Kulakov, L. A. Analysis of transduction in wastewater bacterial populations by targeting the phage-derived 16S rRNA gene sequences. *FEMS Microbiol. Ecol.* **2011**, *76*, 100–108.

Diemer, G. S.; Stedman, K. M. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol. Direct* **2012**, *7(13)*, 13.

Dutilh, B. E.; Thompson, C. C.; Vicente, A. C.; Marin, M. A.; Lee, C.; Silva, G. G.; Schmieder, R.; Andrade, B. G.; Chimetto, L.; Cuevas, D.; Garza, D. R. Comparative genomics of 274 Vibrio cholerae genomes reveals mobile functions structuring three niche dimensions. *BMC Genomics* **2014**, *15(1)*, 1.

Edwards, R.; McNair, K.; Wellington-Oquri, M.; Faust, K.; Raes, J.; Dutilh, B. E. Computational approaches to predict bacteriophage-host relationships, *FEMS Micriobiol. Rev.* **2016**, fuv048.

Faulwetter, J. L.; Burr, M. D.; Parker, A. E.; Stein, O. R.; Camper, A. K. Influence of season and plant species on the abundance and diversity of sulfate reducing bacteria and ammonia oxidizing bacteria in constructed wetland microcosms. *Microb. Ecol.* **2013**, *65(1)*, 111-127.

Fish, J. A.; Chai, B.; Wang, Q.; Sun, Y.; Brown, C. T.; Tiedje, J. M.; Cole, J. R. FunGene: the functional gene pipeline and repository. *Front. Microbiol.* **2013**, *4(4)*, 291.

Fortin, D.; Goulet, R.; Roy, M. Seasonal cycling of Fe and S in a constructed wetland: the role of sulfate-reducing bacteria. *Geomicrobiol. J.* **2000**, *17*, 221–235.

Geider, R. J.; La Roche, J. The role of iron in phytoplankton photosynthesis, and the potential for iron-limitation of primary productivity in the sea. *Photosynth. Res.* **1994**, *39(3)*, 275-301.

Glass, J. B.; Axler, R. P.; Chandra, S.; Goldman, C. R. Molybdenum limitation of microbial nitrogen assimilation in aquatic ecosystems and pure cultures. Beyond the Iron age: the ecological relevance of bioactive trace metals other than Fe and organic growth factors in aquatic systems, **2015**, 7.

Gough, H. L.; Stahl, D. A. Microbial community structures in anoxic freshwater lake sediment along a metal contamination gradient. *ISME J*. **2010**, *5(3)*, 543-558.

Han, Y.; Braatsch, S.; Osterloh, L.; Klug, G. A eukaryotic BLUF domain mediates light-dependent gene expression in the purple bacterium Rhodobacter sphaeroides 2.4.1. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101(33), 12306-12311.

Heidelberg, J. F.; Nelson, W. C.; Schoenfeld, T.; Bhaya, D. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS One* **2009**, *4(1)*, e4169.

Horvath, P.; Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **2010**, *327*,167-170.

Howard, J. B.; Rees, D. C. Structural basis of biological nitrogen fixation. *Chem. Rev*. **1996**, *96(7)*, 2965-2982.

Huson, D. H.; Auch, A. F.; Qi, J.; Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17(3)*, 377-386.

Ivanova, N.; Sikorski, J.; Chertkov, O.; Nolan, M.; Lucas, S.; Hammon, N.; Deshpande, S.; Cheng, J. F.; Tapia, R.; Han, C.; Goodwin, L. Complete genome sequence of the extremely halophilic Halanaerobium praevalens type strain (GSLT). *Stand. Genomic Sci*. **2011**, *4(3)*, 312.

Jackson, E. F.; Jackson, C. R. Viruses in wetland ecosystems. *Freshwater Biol*. **2008**, *53(6)*, 1214-1227.

Jasser, I.; Kostrzewska-Szlakowska, I.; Ejsmont-Karabin, J.; Kalinowska, K.; Weglenska, T. Autotrophic versus heterotrophic production and components of trophic chain in humic lakes: the role of microbial communities. *Pol. J. Ecol.* **2009**, *57(3)*, 423-439.

Jiang, S. C.; Paul, J. H. Gene transfer by transduction in the marine environment. *Appl. Environ. Microbiol.* **1998**, *64*, 2780–2787.

Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; Yamanishi, Y. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*. **2008**, *36(suppl 1)*, D480-D484.

Kenzaka, T.; Tani, K.; Nasu, M. High-frequency phage-mediated gene transfer in freshwater environments determined at single-cell level. *ISME J.* **2010**, *4*, 648–659.

Kerr, Y. H.; Font, J.; Waldteufel, P.; Berger, M. The Soil Moisture and Ocean Salinity Mission-SMOS. *In Geoscience and Remote Sensing Symposium, IGARSS'03 Proceedings IEEE International.* **2000**, *1*, 1-3.

Kidambi, S. P.; Ripp, S.; Miller, R. V. Evidence for phage-mediated gene transfer among Pseudomonas aeruginosa strains on the phylloplane. *Appl. Environ. Microbiol.* **1994**, *60*, 496–500.

Kim, K. H.; Chang, H. W.; Nam, Y. D.; Roh, S. W.; Kim, M. S.; Sung, Y.; Bae, J. W. (2008). Amplification of uncultured single-stranded DNA viruses from rice paddy soil. *Appl. Environ. Microbiol.* **2008**, *74(19)*, 5975-5985.

Kim, O. S.; Imhoff, J. F.; Witzel, K. P.; Junier, P. Distribution of denitrifying bacterial communities in the stratified water column and sediment–water interface in two freshwater lakes and the Baltic Sea. *Aquat. Ecol.* **2011**, *45(1)*, 99-112.

Kochl, S.; Niederstatter, H.; Parson, W. DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. *Forensic DNA Typing Protocols*, **2005**, 13-29.

Koonin, E. V.; Makarova, A. S.; Aravind, L. Horizontal gene transfer in Prokaryotes. Quantification and classification. *Annu. Rev. Microbiol*. **2001**, *55*, 709–742.

Kouki, S.; M'hiri, F.; Saidi, N.; Belaid, S.; Hassen, A. Performances of a constructed wetland treating domestic wastewaters during a macrophytes life cycle. *Desalination* **2009**, *246(1)*, 452-467.

Kristensen, E.; Holmer, M. Decomposition of plant materials in marine sediment exposed to different electron acceptors (O2, NO3 , and SO42), with emphasis on substrate origin, degradation kinetics, and the role of bioturbation. *Geochim. Cosmochim. Acta.* **2001**, *65(3)*, 419-433.

Kunin, V.; He, S.; Warnecke, F.; Peterson, S. B.; Martin, H. G.; Haynes, M.; Ivanova, N.; Blackall, L. L.; Breitbart, M.; Rohwer, F.; McMahon, K. D. (2008). A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res*. **2008**, *18(2)*, 293-297.

Langmead, B.; Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9(4)*, 357-359.

Lasken, R. S.; Stockwell, T. B. Mechanism of chimera formation during the Multiple Displacement Amplification reaction. *BMC Biotechnol*. **2007**, *7(1)*, 1.

Lazar, L.; Gomoiu, M. T.; Boicenco, L.; Vasiliu, D. Total Organic Carbon (Toc) Of the Surface Layer Sediments Covering the Seafloor of the Romanian Black Sea Coast. *GeoEcoMarina* **2012**, *18*, 121.

Leplae, R.; Hebrant, A.; Wodak, S. J.; Toussaint, A. ACLAME: a CLAssification of Mobile genetic Elements. *Nucleic Acids Res.* **2004**, *32(suppl 1)*, D45-D49.

Li, C. H.; Wong, Y. S.; Tam, N. F. Y. Anaerobic biodegradation of polycyclic aromatic hydrocarbons with amendment of iron (III) in mangrove sediment slurry. *Bioresour. Technol.* **2010**, *101(21)*, 8083-8092.

Lima-Mendez, G.; Van Helden, J.; Toussaint, A.; Leplae, R. Prophinder: a computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics* **2008**, *24(6)*, 863-865.

McDaniel, L.; Paul, J. H. Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine Synechococcus species. *Appl. Environ. Microbiol.* **2005**, *71(2)*, 842-850.

Meglitsch, P. Invertebrate Zoology. Oxford University Press. London, England. **1967**.

Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E. M.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; Wilkening, J. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **2008**, *9(1)*, 386.

Middelboe, M. Microbial disease in the sea: effects of viruses on carbon and nutrient cycling (pp. 242-259). Princeton, NJ, USA: Princeton University Press. **2008**.

Mitchell, C. P.; Branfireun, B. A.; Kolka, R. K. Methylmercury dynamics at the upland-peatland interface: Topographic and hydrogeochemical controls. *Water Resour. Res.* **2009**, *45(2)*, W02406.

Mokili, J. L.; Rohwer, F.; Dutilh, B. E. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol*. **2012**, *2*, 63–77.

Motlagh, A. M.; Bhattacharjee, A. S.; Goel, R. Microbiological study of bacteriophage induction in the presence of chemical stress factors in enhanced biological phosphorus removal (EBPR). *Water Res*. **2015**, *81*, 1-14.

Motlagh, A. M.; Bhattacharjee, A. S.; Goel, R. Biofilm control with natural and genetically-modified phages, *World J. Microbiol. Biotechnol*. **2016**, *32*, 1-10.

Motlagh, A.M.; Goel, R. Sustainability of activated sludge processes. Water Reclamation and Sustainability. Waltham: Elsevier Science Ltd. p, 391-414. **2014**.

Olendzenski, L.; Gogarten, J. P. (2009) Evolution of genes and organisms: the tree/web of life in light of horizontal gene transfer. Ann NY Acad Sci. **2009**, *1178*, 137–145.

Overbeek, R.; Begley, T.; Butler, R. M.; Choudhuri, J. V.; Chuang, H. Y.; Cohoon, M.; de Crecy-Lagard, V.; Diaz, N.; Disz, T.; Edwards, R.; Fonstein, M. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* **2005**, *33(17)*, 5691-5702.

Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Parrello, B.; Shukla, M.; Vonstein, V. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res*. **2014**, *42(D1)*, D206-D214.

Paul, J. H. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J*. **2008**, *2(6)*, 579-589.

Pride, D. T.; Wassenaar, T. M.; Ghose, C.; Blaser, M. J. Evidence of host-virus co-evolution in tetranucleotide usage patterns of bacteriophages and eukaryotic viruses. *BMC Genomics* **2006**, *7(1)*, 1.

Puigbo, P.; Wolf, Y. I.; Koonin, E. V. The tree and net components of prokaryote evolution. *Genome Biol. Evol.* **2010**, *2*, 745–756.

Ramasamy, E. V.; Toms, A.; Shylesh, C. M. S.; Jayasooryan, K. K.; Mahesh, M. Mercury fractionation in the sediments of Vembanad wetland, west coast of India. *Environ. Geochem. Health.* **2012**, *34(5)*, 575-586.

Richter, M.; Rossello-Mora, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106(45)*, 19126-31.

Rocha, E. P.; Danchin, A. Base composition bias might result from competition for metabolic resources. *Trends Genet*. **2002**, *18(6)*, 291-294.

Rohwer, F.; Edwards, R. The Phage Proteomic Tree: a genome-based taxonomy for phage. *J. Bacteriol*. **2002**, *184(16)*, 4529-4535.

Rohwer, F.; Thurber, R. V. Viruses manipulate the marine environment. *Nature* **2009**, *459(7244)*, 207-212.

Roossinck, M. J.; Saha, P.; Wiley, G. B.; Quan, J.; White, J. D.; Lai, H.; Chavarria, F.;

Shen, G. A.; Roe, B. A. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Mol. Ecol*. **2010**, *19*, 81-88.

Sano, E.; Carlson, S.; Wegley, L.; Rohwer, F. Movement of viruses between biomes. *Appl. Environ. Microbiol*. **2004**, *70*, 5842–5846.

Seiter, K.; Hensen, C.; Schroter, J.; Zabel, M. Organic carbon content in surface sediments-defining regional provinces, *Deep Sea Research Part I: Oceanographic Research Papers,* **2004**, *51(12)*, 2001-2026.

Short, C. M.; Suttle, C. A. Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Appl. Environ. Microbiol.* **2005**, *71(1)*, 480-486.

Short, S. M.; Suttle, C. A. Use of the polymerase chain reaction and denaturing gradient gel electrophoresis to study diversity in natural virus communities. *Hydrobiologia* **1999**, *401*, 19–32.

Stephens, D. W. Salinity-induced changes in the aquatic ecosystem of Great Salt Lake, Utah. **1998**, 1-8.

Stern, A.; Sorek, R. The phage–host arms race: Shaping the evolution of microbes. *Bioessays* **2011**, *33*, 43–51.

Tatusov, R. L.; Galperin, M. Y.; Natale, D. A.; Koonin, E. V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. **2000**, *28(1)*, 33-36.

Thingstad, T. F.; Bratbak, G.; Heldal, M. Aquatic phage ecology, p. 251-280. In: S. T. Abedon (ed.), Bacteriophage Ecology. Cambridge University Press, Cambridge, UK. **2008**.

Unites States Environmental protection agency, Office of Water. Aquatic Life Ambient Water Quality Criteria for Ammonia - Freshwater, MC 4304T, 820-F-13-013. **2013**.

Wang, H.; Holden, J.; Spera, K.; Xu, X.; Wang, Z.; Luan, J.; Xu, X.; Zhang, Z. Phosphorus fluxes at the sediment–water interface in subtropical wetlands subjected to experimental warming: A microcosm study. *Chemosphere* **2013**, *90(6)*, 1794-1804.

Williams, W.D. Environmental threats to salt lakes and the likely status of inland saline ecosystems in 2025. *Environ. Conserv*. **2002**, *29(02)*, 154-167.

Williamson, S. J.; Rusch, D. B.; Yooseph, S.; Halpern, A. L.; Heidelberg, K. B.; Glass, J. I.; Andrews-Pfannkoch, C.; Fadrosh, D.; Miller, C. S.; Sutton, G.; Frazier, M. The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. *PLoS One* **2008**, *3(1)*, e1456.

Wommack, K. E.; Ravel, J.; Hill, R. T.; Colwell, R. R. Hybridization analysis of Chesapeake Bay virioplankton. *Appl. Environ. Microbiol*. **1999**, *65*, 241–250.

Wurtsbaugh, W. A. Food-web modification by an invertebrate predator in the Great Salt Lake (USA). *Oecologia* **1992**, *89(2)*, 168-175.

Yosef, I.; Manor, M.; Kiro, R.; Qimron, U. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *Proc. Natl. Acad. Sci*. **2015**, *112(23)*, 7267-7272.

Zhang, K.; Martiny, A. C.; Reppas, N. B.; Barry, K. W.; Malek, J.; Chisholm, S. W.; Church, G. M. Sequencing genomes from single cells by polymerase cloning. *Nat. Biotechnol*. **2006**, *24(6)*, 680-686.

Zhou, J.; Bruns, M. A.; Tiedje, J. M. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol*. **1996**, *62(2)*, 316-322.

Zhu, W.; Lomsadze, A.; Borodovsky, M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res*. **2010**, *38(12)*, e132-e132.

CHAPTER 5

THE DIVERSITY OF LYTIC PHAGES AND EFFECTS OF ENVIRONMENTAL STRESS FACTORS ON PROPHAGE INDUCTION OF LYSOGENIC PHAGES IN GREAT SALT LAKE

5.1. Abstract

Bacteriophages are one of the most abundant and significant components of microbial communities in natural ecosystems. Bacteriophages strongly influence the bacterial populations by controlling their size, establishing lysogeny (temperate phages), and introducing genetic exchange via transduction processes. Several studies have been conducted to identify bacteriophages in various ecosystems, but a very small portion of these studies are dedicated to hypersaline water systems and our understanding of the role of phages in the functioning of these important and unique aquatic systems and their role in governing the bacterial diversity remains unclear. Moreover, different environmental stress factors can affect the phage cycle in the ecosystem and induce the prophage on the bacterial chromosome. The objectives of this study were to characterize the lytic phages present in the brine layer sediment of Great Salt Lake and their effects on their bacterial hosts. In addition, the impacts of various physicochemical environmental stresses in the prophage induction process and subsequent bacterial infection on the sediment

samples and the isolated bacterial strains were evaluated. Following isolation of ten different bacterial strains, lytic phages infecting these bacteria were isolated and showed the effects of phage infection on various nutrient cycles. In addition, increase in the pH level and mercury induced the prophages in the isolated bacterial genomes, which corresponds to decrease in the viability of bacterial cells. These results are proving the susceptibility of bacterial hosts in a hypersaline ecosystems due to presence of lytic phages as well as prophage induction.

5.2. Introduction

The importance of bacteriophages in determining the diversity and structure of bacterial community has been a subject of keen interest in freshwater (Middelboe et al., 2009; Kenzaka et al., 2010), terrestrial (Swanson et al., 2009; Wang et al., 2011), and marine (Suttle, 2007; Rowe et al., 2012) ecosystems with several alarming findings (Weinbauer, 2004; Suttle, 2005). The bacteriophages need a host to develop and multiply further, while there are two alternative developmental pathways of phages; lytic cycle causing production of progeny virions and lysogenic cycle resulting in integration of the phage genome into the host chromosome, thus forming lysogens, i.e., host cells bearing integrated phage genomes, called prophages. Under certain conditions, a developmental switch consisting of prophage induction can occur under which damage in host DNA takes place and the induced prophage can enter into lytic cycle (Ptashne, 2004; Wegrzyn and Wegrzyn 2005).

Bacteriophages can help in bacterial evolution process by killing dominant bacteria and helping other minor but phage-resistant bacteria to dominate (Angly et al., 2006). Bacteriophages strongly influence the bacterial populations by controlling their

size, establishing lysogeny (temperate phages), and introducing genetic exchange via transduction processes (Brussow et al., 2004; Reyes et al., 2010; Minot et al., 2011; Modi et al., 2013). However, a very small portion of these studies are dedicated to hypersaline water systems and our understanding of the role of bacteriophages in the functioning of these important and unique aquatic systems and their role in governing the bacterial diversity remains unclear. This research gap in turn questions the environmental sustainability of these ecosystems because the efficient functionality of which directly relates with environmental sustainability.

Viral infection and lysis are among the biotic bottom-up control factors that have significant stimulatory and/or inhibitory effects on biodiversity of prokaryotes, microbial biomass, and their activity in the aquatic system. Viruses can have a relative contribution to bacterial mortality and viruses redirect matter and energy away from higher trophic levels and generate substrate for other noninfected bacteria (Suttle, 2007). Therefore, bacteriophage activity impacts the structure and composition of communities as well as modifying the biogeochemical flux of energy, biomass, and genes in natural ecosystem. Bacteriophages can have cumulatively huge impact not only in abundance, diversity, and composition of host bacteria, but also on nutrient cycling in the natural ecosystems (Suttle, 2005).

Bacteriophages can affect their host through their lytic as well as lysogenic cycles. Lytic phages kill their host immediately. The phage to host ratio is critical in the lytic cycle. The lytic phages do exist with their host but the lytic cycle is not initiated until a suitable phage to host ratio is achieved (Choi et al., 2010; Kotay et al., 2011). In context of natural ecosystems, the phage to host ratio can increase intermittently because

these ecosystems receive discharges from various point and non-point sources. For example, treated municipal effluents contain a high number of total bacteriophages and this can increase the phage to host ratio. The question is whether lytic phages exist in hypersaline ecosystems along with their host bacteria and what phage to host ratio is optimal for infection.

On the other hand, lysogenic cycle of bacteriophages is even more critical because this cycle directly exposes the prokaryotic community in natural ecosystems to environmental stress factors and genetic changes. The temperate phage with lysogenic cycle integrates its genetic element into bacterial DNA and the resulting integrated genetic element is known as the "prophage". When free floating, the phages can act as "carrier" for genetic elements. For example, viral metagenome obtained from marine environment revealed that up to half of the contents of genetic information originated from transduced bacterial and archaeal DNA sequences (Breitbart et al., 2002; Breitbart et al., 2008). The integrated prophage has two implications for their bacterial host; (1) the prophage can stay dormant but can come out of the bacterial DNA (a process known as induction) and get into the lytic cycle and, (2) the prophage can carry certain special genes and if expressed, these gene/s can provide special genetic "identity" to its host bacterium. The first one threatened the survival of the bacterial community and the second one provides more diversity (through phage-mediated horizontal gene transfer) to the prokaryotes in terms of their capabilities to affect the surrounding environment.

Great Salt Lake (Utah, US) as the largest terminal lake in the western hemisphere and one of the world's largest hypersaline lakes confers many environmental, ecological, and economical benefits (Belovsky et al., 2011). Although hypersaline lakes are not

considered common in North America and Europe, saline lakes and inland seas encompass nearly the same water volume as freshwater lakes (Horne and Goldman, 1994). The hypersaline ecosystems are a heterogeneous environment with complex microbial communities which often face several environmental stresses. The biggest question, which is unanswered so far, is whether these environmental stress factors can cause prophage induction in the bacterial communities. Therefore, the objective of this study is characterization of lytic and lysogenic phages in the sediments of Great Salt Lake and understanding the effects of environmental stress factors in prophage induction and overall nutrient removal capacity.

5.3. Materials and methods

5.3.1. Site description and sediment sampling

Great Salt Lake is a terminal lake fed by Bear, Weber, and Jordan rivers. These rivers carry more than 1.1 million tons of salts annually into the lake, which results in one of the most hypersaline water bodies in the world with a salinity gradient of 6% to 28%. The lake is separated by a railroad causeway into North and South Arms. The South Arm receives approx. 95% of the lake's surface inflows and therefore has a lower salinity and contains more diverse biota (Figure 5.1). The sediment sample was collected from 6 mile west of Antelope Island in the South Arm of Great Salt Lake (40°53'51.3" N, 112°21'00" W). Using a box corer (Wildco, FL), sediment samples were collected from deep brine layer in 27-ft depth and stored on ice followed by filtering through 0.45 µm cellulose-acetate membrane syringe filter (Millipore, MA) after transporting to the laboratory.

5.3.2. Quantification of sediment bacterial cell

The sediment bacterial cells were counted by direct plating on different agar media as well as direct staining and enumerating using epifluorescent microscopy (EFM). For direct plating, 10 g of sediment sample was resuspended in PBS solution (8 g/l of NaCl, 0.2 g/l of KCl, 1.44 g/l of Na_2HPO_4, 0.24 g/l of KH_2PO_4, pH of 7.4), vigorously vortexed and allowed to settle down the soil aggregate. The supernatant was serial diluted before inoculation and surface plated onto three different media plates and incubated at 25°C. The media agar plates used in this study were LB (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl), R2A (0.5 g/l proteose peptone, 0.5 g/l casamino acids, 0.5/l yeast extract, 0.5 g/l dextrose, 0.5 g/l starch, 0.3 g/l dipotassium phosphate), and tryptic-soy agar (TSA) (15 g/l tryptone, 5 g/l soytone, 5 g/l sodium chloride). The viable count surface plate method was performed by plating 0.1 ml of different dilutions of sediment sample on the surface of dried agar plates and spread with a sterile inoculating loop. Plates were incubated for 3 to 5 days at 25°C until colonies were visible on the plate. For enumeration of total cell count, the slurry supernatant was filtered with 0.45 μm mixed cellulose hydrophilic filter paper (Millipore, MA) to remove any sediment debris. Bacterial samples for direct microscopic counting were obtained aseptically, fixed with paraformaldehyde with final concentration 4% (w/v) for 30 min at 4°C. The total direct cell count was determined by cell quantification stained with 4′,6-diamidino-2-phenylindole (DAPI) with final concentration of 5 μg ml^{-1} under an epifluorescence microscope (BH-40; Olympus Optical, Tokyo, Japan). Triplicate slides were used for the analysis and a minimum of 10 fields were taken for each slide.

5.3.3. Live and dead bacterial enumeration

The bacterial samples were filtered through a 0.22 μm membrane filter using a vacuum manifold. The bacterial cells captured in the membrane filter were stained for 20 min in dark with a mixture of Propidium iodide and SYTO 9 dyes using Baclight™ bacterial viability kit (Molecular Probes Inc.). The slides containing the stained membrane filters were analyzed under BX51 epifluorescence microscopy (Olympus, Japan) using Cy3 and FITC filters to capture live and dead cells, respectively. Twenty random fields of view were digitally captured at magnification of 1000× with a DPI-71 camera and number of live and dead cells was manually counted and averaged.

5.3.4. Quantification of free phages

For virus-like-particles (VLPs) enumeration, 900 μl of filtrate supernatant sample from the previous step was incubated with 1 unit of DNase 1 (Invitrogen, Canada) and 1× DNase 1 buffer for 15 min at room temperature to remove bacterial DNA from the sample. The reaction was terminated by adding 2.5 mM EDTA and incubating for 15 min at 65°C. Following DNase treatment, 35 μl of 0.5M EDTA solution was added and incubated for 10 min at 65°C to stop the DNase activity. DNase treated sample was vacuum filtered through a stack of 13 mm filters consisting of a 0.02 μm Anodisc membrane filter (Whatman, England), a 0.65 μm Durapore hydrophilic membrane filter (Millipore, MA), and a glass fiber prefilter (Millipore, MA). Anodisc containing captured virus like particles (VLPs) were stained with 1× SYBR Gold dye (Invitrogen, CA), incubated for 20 min in the dark, and analyzed under BX51 epifluorescence microscopy (Olympus, Japan) using FITC filter. VLPs (particles smaller than 0.5 μm) were enumerated from 20 random fields of view that were digitally captured at magnification

of 1000× with a DPI-71 camera. In addition, phage samples were prepared for transmission electron microscopy (TEM) as described in our previous study (Motlagh et al. 2015), and viewed with JEM 1400 transmission electron microscopy (JEOL, Japan).

5.3.5. Bacterial isolation and characterization

The sediment suspension was prepared following the procedure that was mentioned earlier. The supernatant fraction was serial diluted in the PBS solution for subsequent inoculation and surface plated to the isolation media in agar plates. Plates were incubated for 1 to 3 days at 30°C until colonies were visible on the plate. Individual bacterial colonies were plucked and streaked onto another fresh media agar. The re-streaking procedure of the isolated colonies was repeated two more times to ensure that the isolated bacterial colonies are purified followed by storing in 50% (v/v) glycerol at -80ºC for further analysis.

The bacterial isolates were also Gram-stained using the Gram stain kit (BD Diagnostic Systems, NJ) following the manufacture's instruction. Briefly, 10 µl of the cells were spread on the slide with a pipette tip and stained with Crystal violet for 1 min and washed with DI water and air dried. The sample was stained with Gram iodine for another 1 min followed by washing with DI water and air drying. The sample was then exposed to Gram decolorize agent and Gram safranin for 15 and 30 sec-1 min, respectively. Finally, the sample was washed with DI water, air dried, and 10-16 µl of DAPCO was added followed by checking the slide under BH-40 phase contract microscopy (Olympus Optical, Japan). Triplicate slides were used for the analysis and minimum of 10 fields were taken for each slide.

For identification of bacterial isolates, the complete 16S rRNA gene was used.

The bacterial isolates were grown to the exponential phase and their genomic DNA was extracted using UltraClean microbial DNA isolation kit (MoBio, CA) kit following product's instructions. The extracted DNA was used as a template in a PCR reaction using 2× GoTaq Mastermix with universal primers of 8f and 1492r with thermocycler conditions mentioned in Bhattacharjee et al. (2015). The 16S rRNA amplicons were verified on 1% (w/v) agarose gel and purified using QIAquick Gel Extraction Kit (Qiagen, CA).

Prior to identification of isolated bacterial strains by pyro-sequencing, the restriction enzyme were used to identify the replicate bacterial strain and minimize sequencing. Therefore, 5 µl of PCR product was mixed with 1.5 µl of 10x buffer and 0.5 µl of 10× MspI restriction enzyme (NEB, MA) and added to 8 µl of NF water and incubate at 37ºC for 1 hr followed by loading on 1% electrophoresis agarose gel and visualized under UV light. Following screening the replicate strains, the purified amplicons were sequenced on ABI PRISM 377 automated DNA sequencer (Applied Biosystems, USA) in Core Facility at the University of Utah. The DNA sequences were analyzed and identified using BLASTn nucleotide search in NCBI database.

5.3.6. Lytic phage isolation and characterization

The collected sediment samples were used as the initial source for phage isolation. Sediment samples were mixed with 3-fold volume of 1% potassium citrate (10 g/l of potassium citrate, 1.44 g/l of Na_2HPO_4, 0.24 g/l of KH_2PO_4), and vigorously shaked overnight followed by centrifugation at 6000×g for 15 min. The supernatant containing phages was mixed with the equal volume of nutrient broth containing 1 mmol l^{-1} $CaCl_2$ to assist in phage attachment (Ashelford et al., 2003) while 10% (v/v) of

exponentially growing cultures of specific bacterial hosts was added to the solution and incubated at 30ºC for 1 hr. The initial phage enrichment culture was centrifuged at 8000×g for 5 min followed by filtration through 0.45 µm filter before being plated. To this initial phage enrichment culture, 2 ml of 1% (w/v) molten top agar was added, mixed, and plated on 2% (w/v) dried base agar media plate followed by 24 hr incubation at 25ºC. A negative control was also prepared following the same protocol but without any phage added to the culture. Following the overnight incubation, the presence of clear plaque zones was considered evidence of lytic phage susceptibility. The individual phage plaques were plucked and resuspended in SMG buffer (5.8 g/l NaCl, 2 g/l MgSO4-7H2O, 5 ml/l of 5% gelatin solution, 50 ml/l of 1M Tris-HCl pH 7.5) before being replated. The replating of the isolated plaques was repeated two more times to ensure that the isolated phage is purified followed by storing in SMG buffer for further infection.

The phages titres were determined through mixing serial dilutions of isolated phages with the liquid cultures of the corresponding bacterial host strains followed by top agar assay as mentioned above, 24 hr incubation at 30ºC, and enumerating the plaques after the incubation considering the dilution factors. Triplicate plates were prepared for each dilution and plaque numbers were averaged from three plates to calculate the phage titer.

The one-step growth curve was obtained for each of the isolated lytic phages to calculate the latent and eclipse periods, and burst size of the bacteriophages. Therefore, 1 ml of freshly grown bacterial isolate was incubated with the same volume but higher concentration of isolated lytic phage (PBR=10) at 30ºC for 10 min for initial adsorption of viruses. The mixture containing bacterial culture with phage particles was centrifuged

at 3000×g for 5 min followed by discarding the supernatant and resuspending the pellet in fresh bacterial growth medium. The mixture was added to 1 liter freshly grown log-phase bacterial isolate and incubated at 30ºC on shaker. Triplicate samples were taken every 10 min and divided into two aliquots (with and without addition of chloroform) following top agar assay. The samples with adding 1% (v/v) chloroform was used to calculate the total PFU, i.e. free phages and any infectious intracellular phage particles, while the samples without chloroform was used to calculate the free PFU, i.e. extracellular and non-adsorb phage particles. Following the incubation, the PFU counts obtained from triplicate plates were averaged and plotted versus time to obtain the one-step growth curve. The burst size, i.e. amount of infectious phages produced per infected bacterial cell, was calculated by dividing the number of VLPs released from the cell to number of VLPs initially added.

Phage isolates were also analyzed for host range and specificity with performing these assays using the bacterial strains isolated from Great Salt Lake sediment. The soft agar layers containing bacterial lawn of isolated strains were subjected to each of the phage isolates and the specificity of the bacterial strains was detected by the appearance of lytic plaque zone.

5.3.7. Effects of environmental stress factors on bacterial species with prophages in their genome

Environmental impacts to Great Salt Lake are anthropogenic activities and natural processes. Human activities posing a heightened or significant threat to Great Salt Lake include mining, pollution, biological disturbances, anthrapogenically-induced climatic and atmospheric changes (Williams, 2002). As a result of all of these increasing

environmental stress factors, different physical and chemical stresses that are frequently being occurred to Great Salt Lake ecosystem were applied on the sediment as well as the isolated bacterial strains.

In order to evaluate the prophage activation, all of the isolated bacterial hosts were exposed to mitomycin C during bacterial log-phase growth to induce phage transcription and production. Mitomycin C was used as a universal prophage inducer by mixing 10 ml of mitomycin to have a final concentration of 1 µg/ml with equal volume of exponentially grown bacterial culture and incubating overnight on a gyratory shaker followed by phage extraction procedures as mentioned above. The following day, 1% (v/v) of chloroform was added to the solution to release any phages inside the bacterial cells followed by centrifugation at 8000×g for 1 hr and phage extraction process on the supernatant as described above.

The effects of various physico-chemical stress factors including pH, mercury as heavy metal stress factor, and nodularin as cynotoxin widely available in Great Salt Lake were studied on prophage induction and following bacterial viability. The isolated bacteria were grown overnight and duplicate samples of the mid-log phase grown bacteria were spiked with different concentrations of the stress factors. The bacteria were incubated on a shaker at 30°C for 12 hr followed by VLPs enumeration and Baclight assay. In addition, effects of environmental stresses were evaluated on the sediment samples as well by resuspending 1 grams of sediment samples in 10 ml of milliQ water and subjecting the produced slurry mixture to mitomycin C with final concentration of 1 µg/ml as a surrogate for environmental stresses. The slurry was placed on a gyratory shaker for 12 hr followed by phage enumeration and bacterial enumeration.

5.4. Results and discussion

5.4.1. Quantification of sediment bacterial cell

Direct plating of the sediment sample showed counts ranging between 7.2-26.8×10^3 CFU per gram of sediment with four different bacterial colony types which have been recovered on the agar plates. In addition, bacterial population was analyzed using DAPI staining and showed 1.6-3.5×10^6 cells per gram of sediment sample from the Great Salt Lake, while 55±8% of the cells were scored as viable using Live/Dead Baclight^{TM} bacterial viability assay.

5.4.2. Quantification of free phages

Bacteriophage population was analyzed using SYBR Gold staining and showed 2.3-8.5×10^7 virus-like-particles (VLPs) per gram of sediment sample from the Great Salt Lake. This population is a considerable number comparable to or greater than the number of VLPs found in most of the aquatic systems (Suttle, 2007), suggesting that there must be significant phage-bacteria interactions in Great Salt Lake ecosystems. Phenotypic diversity of the phages was examined by electron microscopy and phages were identified using morphological criteria of the International Committee on Taxonomy and Viruses (ICTV) (Ackermann et al., 1992; Pringle, 1998). As shown in Figure 5.2, morphology-based analysis of the viral samples showed that all of the bacteriophages were most similar to the tailed bacteriophage *Caudovirales* with the icosahedral heads with diameters between 15 and 450 nm, and tail length ranges from 35 to 530 nm.

The most abundant phages had icosahedral heads with diameter 50-155 nm. Based on the morphology-based analysis, the phages could be assigned to three virus families including *Myoviridae* with icosahedral head and long contractile tail (Figure

5.2(a) to 5.2(f)), *Siphoviridae* with icosahedral head and long flexible tail (Figure 5.2(g) to 5.2(i)), and *Podoviridae* with icosahedral head and short tail (Figure 5.2(j) to 5.2(l)). The uniquely high proportion of *Siphoviridae* found in the Great Salt Lake sediment suggests that most bacteriophages in this ecosystem are susceptible to prophage induction as a large proportion of temperate phages belong to *Siphoviridae* family (Bench et al., 2007). In addition, as suggested that virulent phages with broad host ranges are members of *Myoviridae* and with narrower host ranges belongs to *Podoviridae* families (Sullivan et al., 2003), the morphology analysis can reveal a snapshot of phage genotypic diversity.

Tailed viruses are the most common isolates infecting prokaryotic hosts residing in hypersaline environments (Sencilo and Roine, 2014). However, the viruses without any observable tail structure could be either bacteriophages that do not have tails in their structure or the bacteriophages that have lost their tails during the preparation procedure particularly in cesium chloride purification step prior to electron microscopy visualization. It should be also explained that all of the virus particles visualized under epifluorescent and electron microscopies might not be bacteriophages and eukaryotic viruses, such as viruses infecting phytoplankton, could be also included in the sampled population. However, since the population of phytoplanktons (algae) was significantly low at the time of sampling event (i.e. end of June) due to high temperature and presence of brine shrimp, which intensely graze the phytoplankton crop, we conclude that the majority of these viruses are likely to be bacteriophages.

5.4.3. Bacterial isolation and characterization

Following direct plating of the sediment sample onto different media agar plates and re-streaking of the colonies, 90 pure colonies were plucked and purified. Using

restriction enzyme followed by PCR on universal 16S rRNA primers, the isolated colonies showed 10 different band pattern on the agarose gel. Pyro-sequencing on one of the strains of each pattern was carried out to identify the bacterial species. Using BLAST search of 16S rRNA sequences against NCBI bacterial database identified 10 different bacterial species, as shown in Table 5.1. *Bacillus pumilus* has a significant role in ecosystem biochemistry with its functions as plant growth promoter and nitrogen fixation transforming atmospheric nitrogen into ammonia (Hernandez et al., 2009). S*taphylococcus epidermidis* as another isolated species is involved in nitrogen cycle with nitrate reduction process.

Although the previous studies showed that S*taphylococcus epidermidis* is a facultative anaerobe (Rowlinson et al. 2006), but the isolated species was cultured in an aerobic condition as it grows better in aerobic environments. *Aeromonas* species including *A. salmonicida* are found ubiquitously in the environment, particularly in fresh or estuarine brackish water (Janda and Abbott, 2010). Although *A. salmonicida* A449 is reported as a facultative species, in this study, it was isolated in an aerobic environment. *Aeromonas* species can be involved in nitrogen cycles and have shown their ability of denitrification and dissimilatory nitrate reduction to nitrous oxide in an anaerobic condition (Hunter and Kuykendall, 2006). *Shewanella putrefaciens* has ubiquitous presence in natural ecosystems with the ability to reduce a wide range of terminal electron acceptors including ferric, nitrate, and nitrite (Dichristina, 1992). Therefore, it plays a key role in the nitrogen cycle as well as metal reduction.

It should be mentioned that the purpose of this section was not to obtain all bacterial isolate present in the sediments. Rather, few of the representative bacterial

communities involved in various nutrient cycles were isolated to evaluate the effects of different environmental stress factors on the prophage induction. The vast majority of the isolated bacterial species were tolerant to high salinities but did not require hypersaline conditions for growth. Gram-staining showed most of the isolated bacteria as Gram-negative isolates, which can be attributed to their tolerance to hypersaline environment.

5.4.4. Lytic phage isolation and characterization

Although a total of 10 bacterial strains were isolated, purified, and characterized from Great Salt Lake sediment, the enrichment and isolation of lytic phages able to infect the representative bacterial isolates was only successful for two of the isolated species including GSL8 and GSL9. The lytic phage corresponding to their bacterial isolates produced clear visible plaques ranging 2-4 mm in diameter. The phage latent and eclipse period and burst size of both phage isolates were calculated based on the one-step growth curve. During the latent period, active virions significantly increases inside the cells but newly synthesized virions do not appear outside the cell as visible plaques. The latent period for phage Φ GSL8 and phage Φ GSL9 were estimated to be 30 min and 20 min, respectively. On the other hand, in the eclipse period, the virus capsid head containing viral nucleic acid separates from its protein coat, resulting in an uncoated virus, and as a result, no infectious virus can be detected during this phase. Based on the one-step growth curve, the eclipse period for phage Φ GSL8 and phage Φ GSL9 were calculated to be 25 min and 15 min, respectively. In addition, burst size for both bacteriophages was calculated and it was estimated to produce 120 and 55 phage particles per infected bacterium.

The infection of GSL8 (*Aeromonas hydrophila* strain ATCC 7966) as one of the

bacterial species involved in nitrogen and sulfate assimilation (Seshadri et al., 2006) proved their susceptibility in phage infection and subsequent effects on nutrient cycle. Stapleton et al. (2005) also showed that *Shewanella putrefaciens* 200 is a metal reducing bacterium and the infection of the GSL9 with the existing phages can affect the iron and magnesium reduction process in the ecosystem. This metal reduction can be coupled with anaerobic respiration, which eventually impacts the organic carbon oxidation.

The host range and specificity analysis using the isolated bacterial strains and lytic phages showed that in all of the assays, phages were able to infect only the host strains used for their isolation. This narrow host range can be identified as a result of an acquired resistance to the co-occurring viruses, as shown in previous studies for both bacteriophages and cyanophages (Borsheim, 1993; Waterbury and Valois, 1993; Alonso et al., 2002). In addition, in contrast to algal viruses, bacterial viruses showed a highly specific characteristic (Motlagh et al., 2016). Previously, it was studied that in saline environments, approximately 73% of marine bacteriophages only lyse their original host bacterium (Wichels et al., 1998).

5.4.5. Effects of environmental stress factors on bacterial species with prophages in their genome

A search for terminase gene against NCBI RefSeq GenBank using BLASTn revealed that *Bacillus pumilus* ATCC 7061 (GSL3), S*taphylococcus epidermidis* ATCC 12228 (GSL4), *Aeromonas salmonicida* A449 (GSL6), and *Shewanella putrefaciens* 200 (GSL9) have 8, 3, 5, and 2 prophages, respectively. Although the prophages on these bacterial genomes were incomplete and/or questionable (except *Shewanella putrefaciens* that has only two intact prophages on its genome), but the presence of prophages (intact,

incomplete, questionable) on the genome of these environmentally relevant organisms indicates that these bacteria might be susceptible to induction under environmental stress factors. Following mitomycin C exposure, significant decrease in bacterial growth (as measured by optical density) and considerable increase in phage population (as measured by fluorescent microscopy as VLPs) was observed, demonstrating the susceptibility of these bacteria to environmental stresses. Mitomycin C causes breaking in the single-strand DNA and activating RecA protein, which results in autoproteolysis of repressors that govern the SOS response and phage lysogeny. In addition, various physical and chemical stresses that are frequently being occurred to Great Salt Lake ecosystem were applied to the isolated bacterial strains.

Water diversion of inflows to Great Salt Lake is increasing due to agriculture, industry, and urbanization. In addition, anthropogenic climate change with increasing the temperature and acid rain precipitation from industrial emissions can fluctuate the pH of the lake. Therefore, pH was varied by titrating the bacterial medium with the addition of HCl or NaOH ranging the final pH between 6 and 10 and its effect on prophage induction and bacterial viability was evaluated (Figure 5.3).

As shown in Figure 5.3, the number of VLPs significantly increased for GSL4 at pH 10 as a result of an unfavorable condition which leads to prophage induction. Consequently the bacterial viability decreased dramatically at pH 10 due to phage infection. In addition, an increase trend in VLPs was observed in GSL4 at pH 5 as well, which was due to prophage induction in the acidic environment. The bacterial isolate GSL4, which was phylogenically close to *Staphylococcus epidermidis* has been reported to have an optimal pH of 7 (Teufel and Gotz, 1993), and the increase in the VLPs

observed in this study can be correlated to the unfavorable condition in the acidic and alkaline conditions, which results in prophage excretion.

As *S. epidermitis* is a non-biofilm forming strain (Zhang et al., 2003) and more susceptible to phage infection due to lack of biofilm exopolysaccharide which can protect against phage infection (Motlagh et al., 2015), it is suggested that the increase in the VLPs was significantly detrimental to the bacterial cells causing cell lysis.

The bacterial isolate GSL6 also showed an increase trend in VLPs at pH 10, which resulted in the subsequent decrease in bacterial viability. The bacterial viability and growth of GSL3 isolate increased with raising the pH, suggesting that the isolated bacterial strain was alkaliphilic in nature, while no prophage induction was observed in any of the pH values. In another study by Gomaa (2013), it was shown that production of protease which subsequently supports the cell growth was increased in higher pH than neutral in *B. pumilus* ATCC 7061, which is phylogenitically close to GSL3 isolate. The bacterial viability in GSL9 was highest in low pH and decreased with increasing the alkalinity, suggesting an acidic condition for optimal growth of the bacterial strain. It was interesting that although the high pH was not favorable for the isolate, the prophage induction did not occur. The increase in the VLPs resulted in a decline of bacterial viability in the corresponding pH.

Metal contamination of the Great Salt Lake is a result of industrial and mining activities in the lake basin. Besides the physical disturbance caused by mining, it can increase the metal pollutant by discharging the mine wastewater. Furthermore, heavy metals leached from mining dumps can increase these toxicant concentrations in the Great Salt Lake aquatic ecosystems (Williams, 2002). Mercury, as one of the major

heavy metals in Great Salt Lake that are driven from industrial activity with increasing concentration (Wurtsbaugh et al., 2012), was chosen to be studied as the stress factor. Different concentration of Hg(II) was applied to the isolated bacteria in their mid-log phase and the effects on prophage induction and bacterial viability was analyzed.

As illustrated in Figure 5.4, a major increase in the VLPs was observed in GSL6 and GSL9 when mercury concentration reached 50 ng/l and 100 ng/l, respectively. The subsequent considerable decrease in the bacterial viability confirms the bacterial infection with these induced phages. As it was discussed in previous section, GSL9 was phylogenetically close to *S. putrefaciens* and resistance to high concentration of Hg(II) has been shown earlier by Jiang et al. (2012). However, the significant increase in the VLPs for GSL9 suggested that the isolated strain is a mutant of the *S. putrefaciens* wild-type, which is susceptible to high concentration of mercury due to prophage induction that ultimately can affect the bacterial viability.

The bacterial isolate GSL3 showed a noticeable increase in VLPs with 50 ng/l of mercury followed by a significant decrease in the bacterial viability. Although this major decrease in the bacterial viability can be attributed to phage induction, it can be also be affected due to high concentration of mercury. Higher concentration of mercury did not show a noticeable increase in VLPs, while it indicated a significant decrease in the bacterial viability suggesting bacterial lysis due to lethal concentration of mercury for the culture. In previous studies, *B. pumilus* has shown resistance to several heavy metals including zinc, nickel, copper, cobalt, and cadmium but not to mercury (Bruins et al., 2000). Therefore, high concentration of mercury and its toxicity can be the major cause of the GSL3 bacterial lysis. No noticeable changes in the bacterial viability and VLPs

were observed for bacterial isolate GSL4 which suggest the resistance of the isolated strain to mercury. Presence of genes associated with resistance to metal ions including mercury in *S. epidermidis* in previous studies (Zhang et al., 2003) confirm the stability of bacterial strain in presence of high concentration of mercury.

Great Salt Lake as a large shallow terminal lake is considered hypereutrophic and extremely rich in nutrients and minerals. Furthermore, considering (oligo) mixing in the Great Salt Lake with salinity enhancement will increase the mixing that eventually results in release of nutrients from brine layer to the photic zone. Nutrient concentrations due to agricultural fertilizers and treated sewage from Jordan River's inflow accumulate in Great Salt Lake as the terminal lake and nutrients are expected to increase in future years. As a result, growth of cyanobacteria can be accelerated and cyanobacterial blooms may cause detrimental effects, including increase of cynotoxin in the Great Salt Lake ecosystems. The genus *Nodularia* majorly exists in sline and coastal lake (Karjalainen, 2005) and therefore, nodularin (NOD), as a cyclic pentapeptide and hepatotoxic cyanobacterial toxins produced by the cyanobacterium *Nodularia spumigena*, was considered as one of the major environmental stress factors for GSL as a terminal hypersaline lake. The effects of nodularin with different concentrations ranging between 1 µg/l to 100 µg/l on prophage induction and bacterial lysis of isolated bacterial strains were studied. Although nodularin even in low concentration of 10 µg/l caused significant decrease in bacterial viability, no prophage induction occurred in any of the isolated bacterial strains. Although nodularin was not able to induce the prophages in the isolated strains, it caused an increase in the VLPs when applied on the sediment sample. Nodularin as one of the microcystin toxins can be part of the cyanobacterial defense mechanism against other

bacteria. Besides the direct effects of nodularin in growth inhibition of other bacteria, it can also induce bacterial prophage, infect the bacterial host, and indirectly affect the bacterial population. In addition, microcystin can change the permeability of bacterial cell wall and consequently increase their sensitivity to other stresses such as phage infection (Dixon et al., 2004).

Exposure of sediment mixed community to mitomycin C showed an average of 12× increase in the VLPs, which proved the induction following incubation with mitomycin C to confirm the existence of prophage. These results demonstrate that prokaryotic communities living in sediments has prophage and were susceptible to induction under environmental stress factors.

5.5. Conclusion

The Great Salt Lake as one of the unique hypersaline ecosystems was studied in this research. The diversity of bacteriophages in the brine layer sediment suggested three virus families of *Myoviridae*, *Siphoviridae*, and *Podoviridae* as the most abundant phages. The uniquely high proportion of *Siphoviridae* found in the Great Salt Lake sediment suggests that most bacteriophages in this ecosystem are susceptible to prophage induction as a large proportion of temperate phages belong to *Siphoviridae* family. Isolation of lytic phages from Great Salt Lake showed the co-existence of phages that can infect bacterial host in the hypersaline environment of this terminal lake. In addition, infection of bacterial hosts involved in nutrient cycling illustrated the influence of phage in bacterial diversity, populations, and indirect effects on various nutrient cycling. Speculating the prophage induction in some of the bacterial hosts due to environmental stresses commonly found in Great Salt Lake was significantly important. The prophage

induction results in an increase in the VLPs population, which were responsible for the decline trend in the bacterial population. These results are proving the susceptibility of bacterial hosts in a hypersaline ecosystems due to presence of lytic phages as well as prophages induction. However, additional sampling is required to determine the seasonal fluctuation of bacteriophage and bacterial population to better understand the role of phage in bacterial population and establish an equilibrium between phage and bacteria.

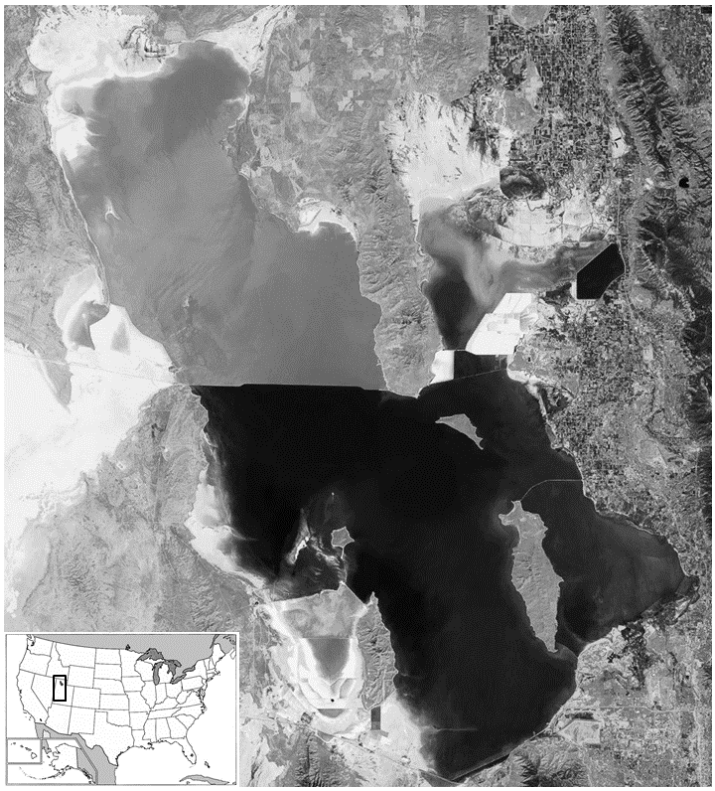

Figure 5.1. The Great Salt Lake as the sampling site of this study is shown in this satellite image. The South Bend of Great Salt Lake is illustrated in dark grey and North Bend with higher salinity in light grey

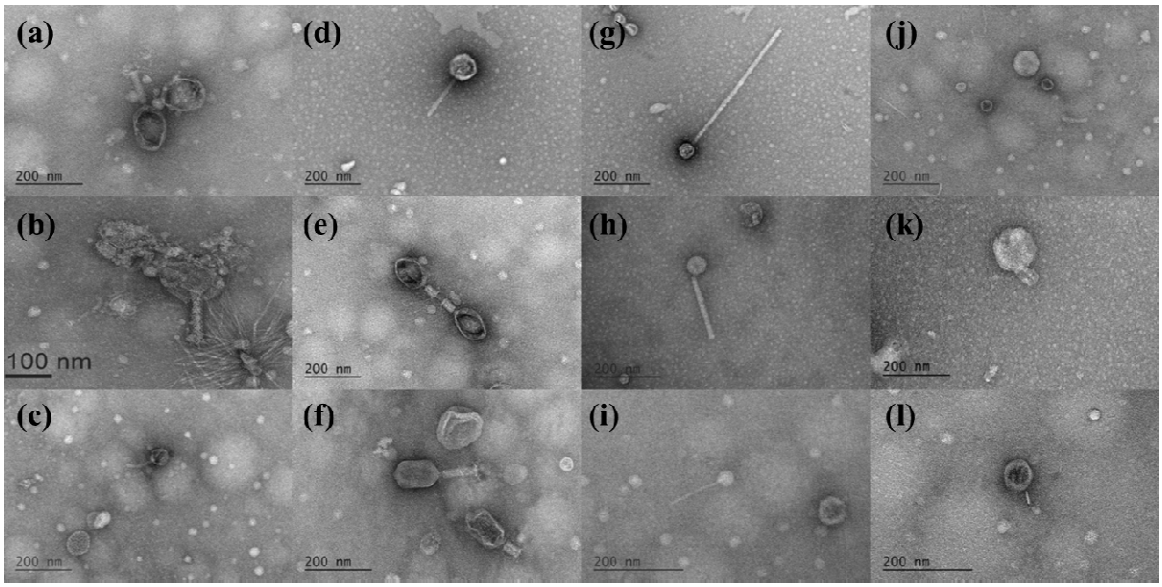

Figure 5.2. *Myoviridae* (micrograph (a) to (f)), *Siphoviridae* (micrograph (g) to (i)), *Podoviridae* (micrograph (j) to (l)) as the most dominant viral families in the sediment

Table 5.1. Isolated bacteria from Great Salt Lake with the matched bacterial strains from NCBI bacterial database

Isolated bacteria	**Gram staining**	**NCBI bacterial database**	**Accession no.**	**avg % ident**
GSL1	+	*Micrococcus yunnanensis* strain TGT-R7	KR476438	99%
GSL2	-	*Sphingomonas aquatilis* strain S7	KF542913	98%
GSL3	-	*Bacillus pumilus* strain ATCC 7061	NR_043242	99%
GSL4	+	*Staphylococcus epidermidis* ATCC 12228	AE015929	99%
GSL5	-	*Sphingomonas echinoides* strain ATCC 14820	NR_024700	99%
GSL6	-	*Aeromonas salmonicida* A449	CP000644	100%
GSL7	+	*Bacillus aquimaris* strain M12	JF411285	100%
GSL8	-	*Aeromonas hydrophila* strain ATCC 7966	NR_074841	99%
GSL9	-	*Shewanella putrefaciens* 200	CP002457	99%
GSL10	-	*Marinobacter lipolyticus*	FN997638	99%

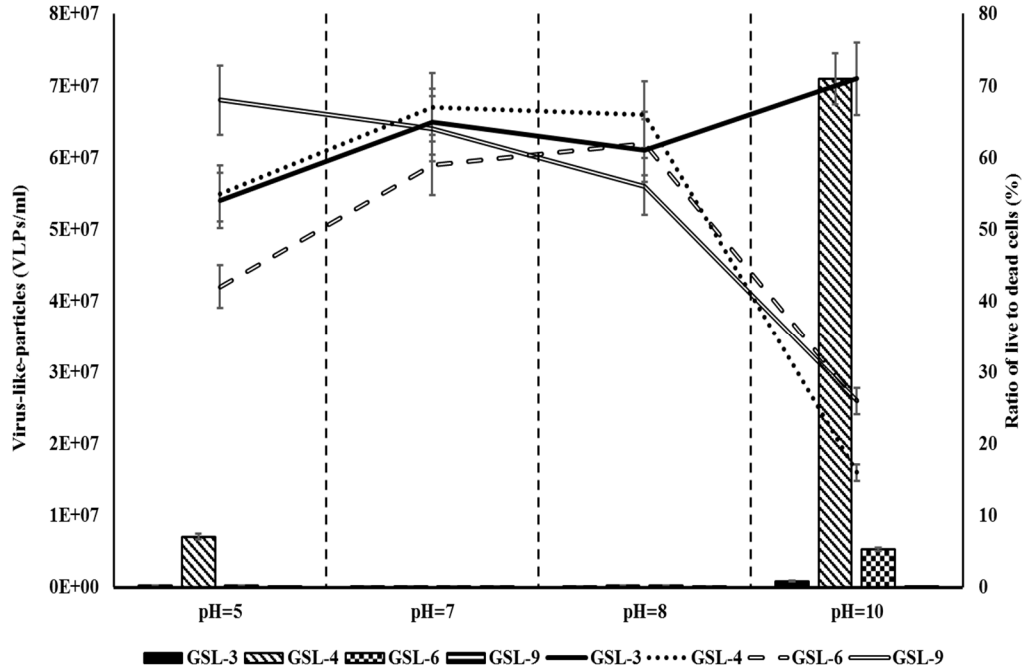

Figure 5.3. Effects of pH fluctuations on prophage induction and bacterial viability of the isolated bacterial strains

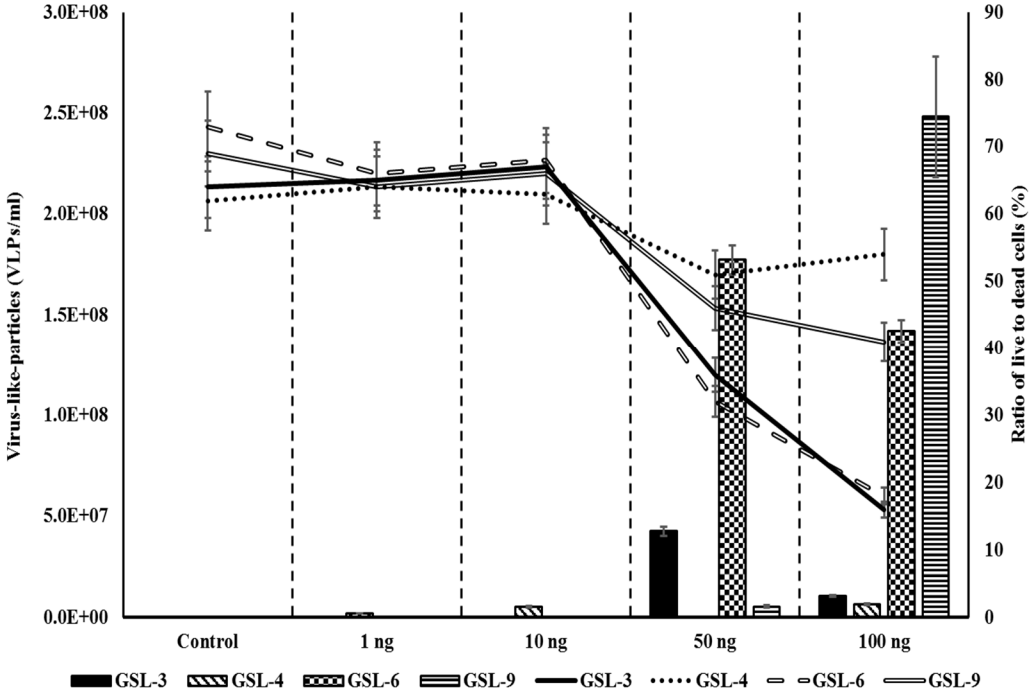

Figure 5.4. Effects of mercury fluctuations on prophage induction and bacterial viability of the isolated bacterial strains

5.6. References

Ackermann, H. W.; DuBow, M. S.; Jarvis, A. W.; Jones, L. A.; Krylov, V. N.; Maniloff, J.; Rocourt, J.; Safferman, R. S.; Schneider, J.; Seldin, L.; Sozzi, T. The species concept and its application to tailed phages. *Arch. Virol*, **1992**, *124(1-2)*, 69-82.

Alonso, M. D. C.; Rodriguez, J.; Borrego, J. J. Characterization of marine bacteriophages isolated from the Alboran Sea (Western Mediterranean). *J. Plankton Res*. **2002**, *24(10)*, 1079-1087.

Angly, F. E.; Felts, B.; Breitbart, M.; Salamon, P.; Edwards, R. A.; Carlson, C.; Chan A.M.; Haynes, M.; Kelley, S.; Liu, H.; Mahaffy, J.M. The marine viromes of four oceanic regions. *PLoS Biol*. **2006**, *4(11)*, e368.

Ashelford, K. E.; Day, M. J.; Fry, J. C. Elevated abundance of bacteriophage infecting bacteria in soil. *Appl. Environ. Microbiol*. **2003**, *69*, 285–289.

Belovsky, G. E.; Stephens, D.; Perschon, C.; Birdsey, P.; Paul, D.; Naftz, D.; Baskin, R.; Larson, C.; Mellison, C.; Luft, J.; Mosley, R. The Great Salt Lake Ecosystem (Utah, USA): long term data and a structural equation approach. *Ecosphere* **2011**, *2(3)*, art33.

Bench, S. R.; Hanson, T. E.; Williamson, K. E.; Ghosh, D.; Radosovich, M.; Wang, K.; Wommack, K. E. Metagenomic characterization of Chesapeake Bay virioplankton. *Appl. Environ. Microbiol*. **2007**, *73(23)*, 7629-7641.

Bhattacharjee, A. S.; Choi, J.; Motlagh, A. M.; Mukherji, S. T.; Goel, R. Bacteriophage therapy for membrane biofouling in membrane bioreactors and antibiotic-resistant bacterial biofilms. *Biotechnol. Bioeng*. **2015**, *112(8)*, 1644-1654.

Borsheim, K. Y. Native marine bacteriophages. *FEMS Microbiol. Ecol*. **1993**, *102*, 141-159.

Breitbart, M.; Haynes, M.; Kelley, S.; Angly, F.; Edwards, R. A.; Felts, B.; Mahaffy, J. M.; Mueller, J.; Nulton, J.; Rayhawk, S.; Rodriguez-Brito, B. Viral diversity and dynamics in an infant gut. *Res. Microbiol*. **2008**, *159(5)*, 367-373.

Breitbart, M.; Miyake, J. H.; Rohwer, F. Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiol. Lett*. **2004**, *236*, 249–256.

Breitbart, M.; Salamon, P.; Andresen, B.; Mahaffy, J. M.; Segall, A. M.; Mead, D.; Rohwer, F. Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci*. **2002**, *99(22)*, 14250-14255.

Bruins, M. R.; Kapil, S.; Oehme, F. W. Microbial resistance to metals in the environment. *Ecotoxicol. Environ. Saf.* **2000**, *45(3)*, 198-207.

Brussow, H.; Canchaya, C.; Hardt, W. D. Phages and the evolution of bacterial

pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev*. **2004**, *68*, 560–602.

Choi, J.; Kotay, S. M.; Goel, R. Various physico-chemical stress factors cause prophage induction in *Nitrosospira multiformis* 25196-an ammonia oxidizing bacteria. *Water Res*. **2010**, *44(15)*, 4550-4558.

Dichristina, T. J. Effects of nitrate and nitrite on dissimilatory iron reduction by *Shewanella putrefaciens* 200. *J. Bacteriol*. **1992**, *174(6)*, 1891-1896.

Dixon, B.; Presley, S. M.; Rainwater, T. R.; Austin, G. P.; Platt, S. G.; Zak, J. C.; Cobb, G. P. Natural disaster microbiology. *Microbe* **2008**, *3*, 312-313.

Gomaa, E. Z. Optimization and characterization of alkaline protease and carboxymethyl-cellulase produced by Bacillus pumillus grown on Ficus nitida wastes. *Braz. J. Microbiol*. **2013**, *44(2)*, 529-537.

Hernandez, J.; de-Bashan, L.; Rodriguez, D.; Rodriguez, Y.; Bashan, Y. Growth promotion of freshwater microalga Chlorella vulgaris by the nitrogen-fixing, plant growth-promoting bacterium Bacillus pumilus from arid zone soils. *Eur. J. Soil Biol*. **2009**, *45*, 88-93.

Horne, A. J.; Goldman, C. R. Limnology (2nd ed., p. 60). New York: McGraw-Hill. **1994**.

Hunter, W. J.; Kuykendall, L. D. Identification and characterization of an Aeromonas salmonicida (syn Haemophilus piscium) strain that reduces selenite to elemental red selenium. *Curr. Microbiol*. **2006**, *52(4)*, 305-309.

Janda, J. M.; Abbott, S. L. The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin. Microbiol. Rev*. **2010**, *23(1)*, 35-73.

Jiang, S.; Ho, C. T.; Lee, J. H.; Van Duong, H.; Han, S.; Hur, H. G. Mercury capture into biogenic amorphous selenium nanospheres produced by mercury resistant Shewanella putrefaciens 200. *Chemosphere* **2012**, *87(6)*, 621-624.

Karjalainen, M. Fate and effects of Nodularia spumigena and its toxin, nodularin, in Baltic Sea planktonic food webs. Finnish Institute of Marine Research. **2005**.

Kenzaka, T.; Tani, K.; Nasu, M. High-frequency phage-mediated gene transfer in freshwater environments determined at single-cell level. *ISME J*. **2010**, *4(5)*, 648-659.

Kotay, S. M.; Datta, T.; Choi, J.; Goel, R. Biocontrol of biomass bulking caused by Haliscomenobacter hydrossis using a newly isolated lytic bacteriophage. *Water Res.* **2011**, *45(2)*, 694-704.

Middelboe, M.; Holmfeldt, K.; Riemann, L.; Nybroe, O.; Haaber, J. Bacteriophages drive

strain diversification in a marine Flavobacterium: implications for phage resistance and physiological properties. *Environ. Microbiol*. **2009**, *11*, 1971–1982.

Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S. A.; Wu, G. D.; Bushman, F. D. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res*. **2011**, *21(10)*, 1616-1625.

Modi, S. R.; Lee, H. H.; Spina, C. S.; Collins, J. J. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **2013**, *499(7457)*, 219-222.

Motlagh, A. M.; Bhattacharjee, A. S.; Goel, R. Microbiological study of bacteriophage induction in the presence of chemical stress factors in enhanced biological phosphorus removal (EBPR). *Water Res*. **2015**, *81*, 1-14.

Motlagh, A. M., Bhattacharjee, A. S., Goel, R. Bacteriophage therapy for biofilm control with natural and genetically-modified lytic phages. *World J. Microbiol. Biotechnol.* **2016**, *32(4)*, 67-77.

Pringle, C. R. Virus taxonomy. *Arch. Virol*. **1998**, *143(7)*, 1449-1459.

Ptashne, M. A genetic switch: phage lambda revisited (Vol. 3). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. **2004**.

Reyes, A.; Haynes, M.; Hanson, N.; Angly, F. E.; Heath, A. C.; Rohwer, F.; Gordon, J. I. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **2010**, *466(7304)*, 334-338.

Rowe, J. M.; DeBruyn, J. M.; Poorvin, L.; LeCleir, G. R.; Johnson, Z. I.; Zinser, E. R.; Wilhelm, S. W. Viral and bacterial abundance and production in the Western Pacific Ocean and the relation to other oceanic realms. *FEMS Microbiol. Ecol*. **2012**, *79(2)*, 359-370.

Rowlinson, M. C.; LeBourgeois, P.; Ward, K.; Song, Y.; Finegold, S. M.; Bruckner, D. A. Isolation of a strictly anaerobic strain of Staphylococcus epidermidis. *J. Clin. Microbiol*. **2006**, *44(3)*, 857-860.

Sencilo, A.; Roine, E. A glimpse of the genomic diversity of haloarchaeal tailed viruses. In The Proceedings from Halophiles, the International Congress on Halophilic Microorganisms (p. 135). *Frontiers Media SA*. **2014**.

Seshadri, R.; Joseph, S. W.; Chopra, A. K.; Sha, J.; Shaw, J.; Graf, J.; Haft, D.; Wu, M.; Ren, Q.; Rosovitz, M. J.; Madupu, R. Genome sequence of Aeromonas hydrophila ATCC 7966T: jack of all trades. *J. Bacteriol*. **2006**, *188(23)*, 8272-8282.

Stapleton Jr, R. D.; Sabree, Z. L.; Palumbo, A. V.; Moyer, C. L.; Devol, A. H.; Roh, Y.;

Zhou, J. Z. Metal reduction at cold temperatures by Shewanella isolates from various marine environments. *Aquat. Microb. Ecol*. **2005**, *38(1)*, 81.

Sullivan, M. B.; Waterbury, J. B.; Chisholm, S. W. Cyanophages infecting the oceanic cyanobacterium Prochlorococcus. *Nature* **2003**, *424(6952)*, 1047-1051.

Suttle, C. A. Viruses in the sea. *Nature* **2005**, *437*, 356–361.

Suttle, C. A. Marine viruses-major players in the global ecosystem. *Nature Rev. Microbiol*. **2007**, *5*, 801–812.

Swanson, M. M.; Fraser, G.; Daniell, T. J.; Torrance, L.; Gregory, P. J.; Taliansky, M. (2009). Viruses in soils: morphological diversity and abundance in the rhizosphere. *Ann. Appl. Biol*. **2009**, *155(1)*, 51-60.

Teufel, P.; Gotz, F. Characterization of an extracellular metalloprotease with elastase activity from Staphylococcus epidermidis. *J. Bacteriol*. **1993**, *175(13)*, 4218-4224.

Wang, G.; Asakawa, S.; Kimura, M. (2011) Spatial and temporal changes of cyanophage communities in paddy field soils as revealed by the capsid assembly protein gene g20. *FEMS Microbiol. Ecol*. **2011**, *76(2)*, 352-359.

Waterbury, J. B.; Valois, F. W. Resistance to co-occurring phages enables marine Synechococcus communities to coexist with cyanophages abundant in seawater. *Appl. Environ. Microbiol*. **1993**, *59(10)*, 3393-3399.

Wegrzyn, G.; Wegrzyn, A. (2005). Genetic switches during bacteriophage lambda development. *Prog Nucleic Acid Res Mol Biol*. **2005**, *79*, 1.

Weinbauer, M. G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev*. **2004**, *28(2)*, 127-181.

Wichels, A.; Biel, S. S.; Gelderblom, H. R.; Brinkhoff, T.; Muyzer, G.; Schutt, C. Bacteriophage diversity in the North Sea. *Appl. Environ. Microbiol*. **1998**, *64*, 4128-4133.

Williams, W. D. Environmental threats to salt lakes and the likely status of inland saline ecosystems in 2025. *Environ. Conserv*. **2002**, *29(02)*, 154-167.

Wurtsbaugh, W. A.; Marcarelli, A. M.; Boyer, G. L. Eutrophication and metal concentrations in three bays of the Great Salt Lake (USA). Final report to the Utah Division of Water Quality, Salt Lake City, Utah. 70 p. **2012**.

Zhang, Y. Q.; Ren, S. X.; Li, H. L.; Wang, Y. X.; Fu, G.; Yang, J.; Qin, Z. Q.; Miao, Y. G.; Wang, W. Y.; Chen, R. S.; Shen, Y. (2003). Genome-based analysis of virulence genes in a non-biofilm-forming Staphylococcus epidermidis strain (ATCC 12228). *Mol. Microbiol*. **2003**, *49(6)*, 1577-1593.

CHAPTER 6

CONCLUSION

This study focused on the role of bacteriophage in bacterial diversity, population, and how they influence the resieliency of the ecosystems. In this research, with a systematic approach, initially the lab-scale and engineered ecosystem was studied. In the lab-scale biological phosphorus removal, there were many factors in-control and can be manipulated to better understand the phage-host interaction and their effects on the bacterial population and metabolic activity of phsophrus accumulation organisms (PAOs). The challenge in studying the effect of chemical stress factors on PAOs was the inability to culture these unique organisms, although they can be enriched in the lab up to high concentrations. Under this challenge, apparently, it was difficult to conclude that all the VLPs came from PAOs. However, based on qPCR results and DAPI staining, we can comfortably conclude that the reactor was enriched in PAOs and PAOs contributed to VLPs number during each chemical stress test. Based on the presented results, 0.5 ppm of copper, 500 µg/l of KCN, and 0.1 µg/ml of ciprofloxacin resulted in the highest prophage induction and coincided with significant decreases in *ppk1* genes and phosphorus removal performances. Prophages harbored in PAOs were induced when exposed to environmental stress factors including heavy metal, toxic chemical, and

antibiotic. Prophage induction in PAOs could be one of the underlying causes for deterioration in phosphorus removal efficiency and EBPR performance under stress. It is well known that wastewater treatment plants are long-term processes, and the limited short-term study of stress factors on PAOs and the subsequent dynamics of bacteriophages cannot reflect all of the influences on biological phosphorus removal. However, it can reveal a greater extent of knowledge about the process of PAOs infection and the deterioration of EBPR following different physical and chemical stress factors. A mixture of free floating phages was isolated from a full-scale EBPR plant and the contents of the reactors were infected with this mixture. It was observed that the dispersed biomass was more sensitive to the added phages and demonstrated that, perhaps, some of the added phages were proved to be lytic to PAOs. These results also have tremendous practical applications. For example, these results emphasize the importance of granular activated sludge processes to diminish the effects of external microbial perturbations such as the induction of prophage and lytic phages infection on key microbial community.

In continuation of the effort to understand the phage-host interactions and their role in affecting various nutrients cycling, a larger scale natural ecosystem was chosen to be studied. Therefore, Great Salt Lake (GSL) as one of the unique hypersaline ecosystems was studied in this research. As one of the most-interesting and less-studied hypersaline environments on the planet, Great Salt Lake showed a very high diversity in bacteria and bacteriophage species, which was related to various nutrient cycles. Metagenomics analyses revealed the relation between the phage and bacteria as their host through their GC contents with respect to the contig coverage, scaffold length, and k-mer

abundance. Furthermore, the presence of prophage and their host in the sediment showed the susceptibility of Great Salt Lake environment in various nutrient cycles. As prophages are able to be induced in presence of different environmental stress factors and can infect their bacterial hosts, which results into alteration of nutrient cycles of the ecosystem. In addition, gene prediction in the bacterial and viral contigs showed the shared functional genes in the samples, which could have originated from gene transfer process between the bacterial host and phage. Lastly, identification of CRISPRs in the bacterial contigs confirms prevalence of previous infections among the bacterial community by phages originated from marine environments.

The diversity of bacteriophages in the brine layer sediment suggested three virus families of *Myoviridae*, *Siphoviridae*, and *Podoviridae* as the most abundant phages. The uniquely high proportion of *Siphoviridae* found in the Great Salt Lake sediment suggests that most bacteriophages in this ecosystem are susceptible to prophage induction as a large proportion of temperate phages belong to *Siphoviridae* family. Isolation of lytic phages from Great Salt Lake showed the co-existence of phages that can infect bacterial host in the hypersaline environment of this terminal lake. In addition, infection of bacterial hosts involved in nutrient cycling illustrated the influence of phage in bacterial diversity, populations, and indirect effects on various nutrient cycling. Speculating the prophage induction in some of the bacterial hosts due to environmental stresses commonly found in Great Salt Lake was significantly important. The prophage induction results in an increase in the VLPs population, which was responsible for the declining trend in the bacterial population.

6.1. Recommendation for future works

The results in this project showed the susceptibility of bacterial hosts in engineered and hypersaline ecosystem due to presence of lytic phages as well as prophages induction. This bacterial susceptibility was revealed to affect the bacterial population, diversity, and eventually nutrient cycling. However, additional studies in other nutrient removing bioreactors are recommended to better understand the role of phages and effects of prophage induction in various nutrient cycling. Furthermore, additional sampling is required to determine the seasonal fluctuation of bacteriophage and bacterial population to better understand the role of phage in bacterial population and establish an equilibrium between phage and bacteria. In addition, there are multiple phages that interact and infect a single bacterial host and understanding the phage-phage interaction network between these various phages capable of infecting a single bacterium is important. Although viral coinfection and phage-phage interactions occurs frequently in nature, the studies related to phage-phage interaction in natural ecosystems particularly hypersaline environments are very limited.